



Université d'Orléans

MASTER ECONOMETRIE ET STATISTIQUE APPLIQUEE (ESA)

Université d'Orléans

Econométrie pour la Finance

Modèles ARCH - GARCH Applications à la VaR

Christophe Hurlin

Documents et Supports

Année Universitaire 2006-2007

Master Econométrie et Statistique Appliquée (ESA)
Université d'Orléans
Faculté de Droit, d'Economie et de Gestion
Bureau A 224
Rue de Blois – BP 6739
45067 Orléans Cedex 2
www.univ-orleans.fr/deg/masters/ESA/

October 28, 2004

Contents

1	Introduction	1
2	Processus linéaires et processus non linéaires	1
2.1	Les principales propriétés des séries financières	1
2.2	Les grandes classes de modèles non linéaires	11
2.2.1	Modèles bilinéaires (Granger et Andersen, 1978)	12
2.2.2	Modèles auto-régressifs exponentiels (modèles EXPAR)	14
2.2.3	Modèles autorégressifs à seuil (modèles TAR)	14
2.3	L'approche ARCH / GARCH et la modélisation de l'incertitude	15
3	Modèles ARCH / GARCH linéaires	17
3.1	Modèles ARCH(q)	17
3.2	Modèle avec erreurs ARCH(q)	21
3.3	Modèles GARCH(p, q)	23
4	Estimation et Prévisions	27
4.1	Estimateurs du MV sous l'hypothèse de normalité et Estimateurs du PMV	27
4.1.1	Maximum et Pseudo Maximum de Vraisemblance appliqués aux modèle ARCH / GARCH	27
4.1.2	La procédure AUTOREG : estimation par MV et PMV	30
4.1.3	La procédure AUTOREG : variances conditionnelles estimées et résidus	34
4.1.4	La procédure MODEL	38
4.2	Estimateurs du MV sous d'autres lois	43
4.2.1	La distribution de Student	43
4.2.2	La distribution de Student dissymétrique standardisée	45
4.2.3	La distribution Generalized Error Distribution	45
4.2.4	La procédure AUTOREG	46
4.2.5	La procédure MODEL	47
4.3	Prévisions et intervalles de confiance	50
4.4	Tests d'effets ARCH / GARCH	50
5	Extension des Modèles ARCH / GARCH linéaires	52
5.1	Application : Value at Risk	52
5.2	Modèles ARMA-GARCH	53
5.3	Modèles GARCH-M	54
5.4	Modèles IGARCH	55
6	Modèles ARCH / GARCH asymétriques	59
6.1	Modèle EGARCH	59

6.2	Modèle GJR-GARCH	62
6.3	Généralisations APARCH et VSGARCH	63
6.4	Modèles TARCH et TGARCH	66
6.5	Modèle QGARCH	68
6.6	Modèles LSTGARCH et ANSTGARCH	69
7	Modèles ARCH et mémoire longue	72
7.1	Modèle FIGARCH	72
7.2	Modèle HYGARCH	73
7.3	Modèle FAPARCH	74
8	Modèles Multivariés	74
9	Conclusion	74

1. Introduction

2. Processus linéaires et processus non linéaires

L'apparition des modèles ARCH / GARCH doit être replacé dans le contexte plus vaste du débat sur la représentation linéaire ou non linéaire des processus stochastiques temporels.

"A major contribution of the ARCH literature is the finding that apparent changes in the volatility of economic time series may be predictable and result from a specific type of nonlinear dependence rather than exogenous structural changes in variables." (Berra et Higgins, 1993, page 315).

Comme l'indiquent Berra et Higgins, la modélisation ARCH / GARCH et ses extensions correspond à une (i) représentation spécifique de la non linéarité (ii) qui permet une modélisation simple de l'incertitude. Nous allons successivement évoquer ces deux points. Mais avant cela passons en revue les principales propriétés des séries financières de prix (action, obligation, taux de change..) et de rendement¹. Ce qui nous permettra au passage d'introduire un certain nombre de définitions essentielles.

2.1. Les principales propriétés des séries financières

Les séries de prix d'actif et de rendements présentent généralement un certain nombre de propriétés similaires suivant leur périodicité. Soit p_t le prix d'un actif à la date t et r_t le logarithme du rendement correspondant:

$$r_t = \log(p_t) - \log(p_{t-1}) = \log(1 + R_t) \quad (2.1)$$

où $R_t = (p_t - p_{t-1})/p_{t-1}$ désigne la variation relative des prix. Considérons à titre d'exemple l'indice Standard & Poor observé en clôture sur la période du 03/07/1989 au 24/11/2003 ainsi que le rendement quotidien associé (figure 2.2 et 2.3). Sous Sas, pour visualiser ces deux séries, on utilise le programme suivant (**fichier example1.sas**) :

Charpentier (2002) distingue ainsi 8 principales propriétés que nous allons successivement aborder.

Propriété 1 (*Stationnarité*) *Les processus stochastiques p_t associés aux prix d'actif sont généralement non stationnaires au sens de la stationnarité du second ordre, tandis que les processus associés aux rendements sont compatibles avec la propriété de stationnarité au second ordre.*

¹Voir Cuthbertson (2000), "Economie Financière Quantitative: Actions, Obligations et Taux de Change", De Boeck, pour la définition des concepts financiers de base.

Figure 2.1: Programme Example1.sas

```

data donnees;
  infile 'C:\Chris\Cours\Econometrie_Finance\AppliSAS\bonds2.csv' dlm=';' ;
  input tc t1 cac sp;
  lsp = log(sp);
  dlsp=dif(lsp);
  t+1;
  keep dlsp lsp t;
run;

proc gplot data=donnees;
  plot dlcac*t=1 / vref=0.08;
  plot dlsp*t=1 / vref=0;
  plot lsp*t=1 / vref=0;
  symbol1 i=join v=none;
run;

quit;

```

Rappelons au passage les définitions de la stationnarité forte et de la stationnarité faible (ou stationnarité du second ordre). Soit un processus temporel aléatoire $(x_t, t \in \mathbb{Z})$.

Definition 2.1. *Le processus x_t est dit strictement ou fortement stationnaire si \forall le n -uplet du temps $t_1 < t_2 < \dots < t_n$, tel que $t_i \in \mathbb{Z}$ et pour tout temps $h \in \mathbb{Z}$ avec $t_i + h \in \mathbb{Z}, \forall i, i = 1, \dots, n$, la suite $(x_{t_1+h}, \dots, x_{t_n+h})$ à la même loi de probabilité que la suite $(x_{t_1}, \dots, x_{t_n})$.*

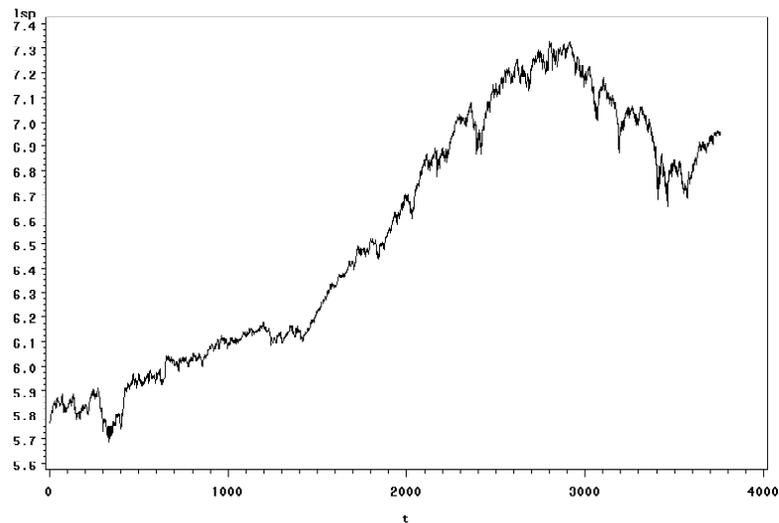
Dans la pratique, on se limite généralement à requérir la stationnarité du second ordre (ou *stationnarité faible*) du processus étudié.

Definition 2.2. *Un processus $(x_t, t \in \mathbb{Z})$ est dit stationnaire au second ordre, ou stationnaire au sens faible, ou stationnaire d'ordre deux si les trois conditions suivantes sont satisfaites :*

- (i) $\forall t \in \mathbb{Z}, E(x_t^2) < \infty$
- (ii) $\forall t \in \mathbb{Z}, E(x_t) = m$, indépendant de t
- (iii) $\forall (t, h) \in \mathbb{Z}^2, cov(x_t, x_{t+h}) = E[(x_{t+h} - m)(x_t - m)] = \gamma(h)$, indépendant de t

La première condition $E(x_t^2) < \infty$ garantit tout simplement l'existence (ou la *convergence*) des moments d'ordre deux. La seconde condition $E(x_t) = m, \forall t \in \mathbb{Z}$ porte sur les moments d'ordre un et signifie tout simplement que les variables aléatoires x_t doivent avoir la même espérance quelle que soit la date t . Autrement dit, *l'espérance du processus x_t doit être indépendante du temps*. Enfin, la troisième condition, $\gamma(h)$

Figure 2.2: Indice SP500 : 03/07/1989 au 24/11/2003



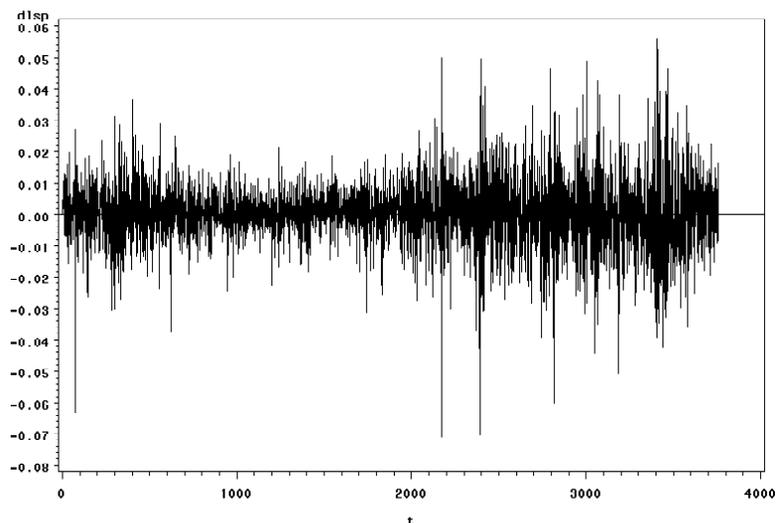
indépendant de t , porte sur les moments d'ordre deux résumés par la fonction d'autocovariance. Cette condition implique que ces moments doivent être indépendants de la date considérée et ne doivent dépendre uniquement que de l'ordre des retards. Autrement dit *la fonction d'autocovariance du processus x_t doit être indépendante du temps*. En résumé, un processus est stationnaire au second ordre si l'ensemble de ses moments sont indépendants du temps. Par conséquent, il convient de noter que la stationnarité implique que la variance $\gamma(0)$ du processus x_t est constante au cours du temps. Sans pratiquer de test de l'hypothèse de non stationnarité ou de stationnarité, on peut observer sur la figure (??) que la dynamique de l'indice SP500 semble ne pas satisfaire aux différents éléments de la définition de la stationnarité du second ordre. Le diagnostic quant à la stationnarité des rendements est plus difficile à prononcer et nécessiterait l'application de tests de l'hypothèse de non stationnarité (ADF, ERS, Max-ADF etc.).

Propriété 2 (*Autocorrélations des carrés des variations de prix*) *La série r_t^2 associée aux carrés des rendements présente généralement de fortes auto-corrélations tandis que les auto-corrélations de la série r_t sont souvent très faibles (hypothèse de bruit blanc).*

L'absence d'auto-corrélation des rendements renvoie à la notion d'efficience. Nous n'entrerons pas ici dans le détail de la théorie et des tests de l'hypothèse de marchés efficients (*Efficient Market Hypothesis* ou *EMH*)². Retenons simplement que sous l'*EMH*,

²Voir Cuthbertson (2000), "Economie Financière Quantitative: Actions, Obligations et Taux de

Figure 2.3: Rendement SP500 : 03/07/1989 au 24/11/2003



le cours p_t d'une action incorpore toutes les informations pertinentes. Ce que l'on entend par information renvoie ici aux différentes formes, faibles, semi-forte et forte de l'efficience. Dis autrement, l'hypothèse d'efficience du marché implique en effet que les rendements anticipés d'équilibre corrigés du risque ne sont pas prévisibles.

Quelle que soit la définition retenue, les cours ne peuvent varier entre t et $t + 1$ qu'en raison de l'arrivée de "nouvelles" (*news*) non anticipées. Sous l'hypothèse d'anticipations rationnelles, les erreurs de prévisions définies par $\varepsilon_{t+1} = P_{t+1} - E_t P_{t+1}$ doivent être nulles en moyenne et ne doivent être corrélées avec aucune information de l'ensemble Γ_t d'information disponible à la date t . Cette dernière propriété est appelée "propriété d'orthogonalité". Or, il est possible de démontrer que si ε_t est auto-corrélé alors la propriété d'orthogonalité n'est pas respectée. Par exemple, supposons ε_t suit un processus $AR(1)$, $\varepsilon_{t+1} = \rho\varepsilon_t + v_t$ où v_t désigne un bruit blanc. L'erreur de prévision $P_{t+1} - E_t P_{t+1}$ ou *profit non anticipé*, est connu en partie à la date t et par conséquent forme une partie de Γ_t , i.e. $E(\varepsilon_{t+1}/\Gamma_t) \neq 0$. Dis autrement, l'erreur de prévision à la date t améliore les prévisions à la période suivante et donc aide à prévoir le cours de la date $t + 1$.

On applique souvent l'hypothèse de marchés efficients aux rendements des actions r_t . L'efficience informationnelle implique alors que personne ne peut dégager de profit anormal en achetant et en revendant une cation. Ainsi, on retrouve la même équation que précédemment pour les rendements:

$$r_{t+1} = E_t r_{t+1} + \varepsilon_{t+1} \quad (2.2)$$

Change", De Boeck, chapitres 5 et 6.

où l'erreur de prévision ε_{t+1} vérifie $E(\varepsilon_{t+1}) = 0$. Les rendements réels vont être tantôt inférieurs, tantôt supérieurs aux rendements réels, mais en moyenne les rendements non anticipés ε_{t+1} sont nuls. Sous l'hypothèse d'anticipation rationnelle, rendements non anticipés vérifient en outre une condition d'orthogonalité par rapport à un ensemble d'information Ω_t . Tout comme pour le cours, *la présence d'une auto-corrélation des rendements r_t constitue une violation de l'hypothèse d'EMH sous l'hypothèse d'anticipation rationnelle. En revanche, on notera que l'hypothèse EMH n'impose a priori aucune restriction sur la forme des moments supérieurs à un de la distribution de ε_t . Par exemple, la variance de ε_{t+1} peut être liée avec ses valeurs passées tout en respectant l'efficacité informationnelle. L'hypothèse d'anticipation rationnelle n'impose des restrictions que sur le premier moment de ε_t . Dès lors, l'auto-corrélation des r_t^2 n'est pas incompatible avec l'EMH.*

On vérifie sur nos données relatives à l'indice SP500 l'absence de corrélations des rendements ainsi que la présence de corrélations des rendements aux carrés r^2 . Pour cela, on considère le statistique de Box et Pierce. On note a_k l'autocorrélation d'ordre k du processus $\{z_t, t \in \mathbb{Z}\}$. Pour un ordre K , le test de Box et Pierce est le test de l'hypothèse $H_0 : a_1 = \dots = a_K = 0$ contre $H_1 : \exists j \in [1, K]$, tel que $a_j \neq 0$. La statistique de ce test est :

$$Q_{BP} = T \sum_{k=1}^K \hat{a}_k^2 \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \chi^2(K) \quad (2.3)$$

où \hat{a}_k désigne l'autocorrélation empirique :

$$\hat{a}_k = \frac{\sum_{t=k+1}^T (z_t - \bar{z})(z_{t-k} - \bar{z}) / (T - k)}{\sum_{t=k+1}^T (z_t - \bar{z})^2 / T}$$

De la même façon la statistique de Ljung-Box, définie pour un ordre K , correspond à l'hypothèse nulle $H_0 : a_k = 0, \forall k \leq K$ et vérifie :

$$Q_K = T(T+2) \sum_{k=1}^K \frac{\hat{a}_k^2}{T-k} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \chi^2(K) \quad (2.4)$$

Sur la figure ??, l'hypothèse nulle d'absence d'autocorrélations des rendements r_t est acceptée jusqu'à l'ordre maximal testé $K = 20$, tandis que cette même hypothèse est rejetée en ce qui concerne le carré des rendements à l'ordre $K = 1$ (figure ??)

Correlogram of R_SP						
Date: 10/01/04 Time: 16:36						
Sample: 1 3756						
Included observations: 3755						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.005	-0.005	0.1112	0.739
		2	-0.015	-0.015	1.0087	0.604
		3	-0.037	-0.037	6.2220	0.101
		4	-0.014	-0.014	6.9153	0.140
		5	-0.019	-0.020	8.2728	0.142
		6	-0.025	-0.028	10.701	0.098
		7	-0.038	-0.040	16.093	0.024
		8	0.004	0.001	16.154	0.040
		9	0.006	0.002	16.294	0.061
		10	0.025	0.021	18.570	0.046
		11	-0.028	-0.029	21.479	0.029
		12	0.029	0.027	24.575	0.017
		13	0.039	0.039	30.447	0.004
		14	0.011	0.010	30.877	0.006
		15	0.007	0.011	31.051	0.009
		16	0.002	0.007	31.065	0.013
		17	-0.023	-0.019	33.000	0.011
		18	-0.013	-0.012	33.634	0.014
		19	0.002	0.006	33.646	0.020
		20	-0.036	-0.035	38.499	0.008

Corrélogramme des Rendements du SP500

Correlogram of R_SP2						
Date: 10/01/04 Time: 16:36						
Sample: 1 3756						
Included observations: 3755						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.191	0.191	137.22	0.000
		2	0.173	0.141	249.19	0.000
		3	0.162	0.113	348.33	0.000
		4	0.114	0.050	397.09	0.000
		5	0.139	0.082	470.00	0.000
		6	0.147	0.084	551.81	0.000
		7	0.130	0.058	615.71	0.000
		8	0.123	0.046	672.88	0.000
		9	0.124	0.048	730.42	0.000
		10	0.127	0.052	791.49	0.000
		11	0.121	0.041	846.41	0.000
		12	0.119	0.037	899.35	0.000
		13	0.094	0.010	932.55	0.000
		14	0.095	0.017	966.25	0.000
		15	0.079	0.002	990.07	0.000
		16	0.113	0.046	1038.6	0.000
		17	0.094	0.017	1071.8	0.000
		18	0.110	0.037	1117.2	0.000
		19	0.115	0.039	1167.4	0.000
		20	0.083	0.003	1193.3	0.000

Corrélogramme des Rendements au Carré du SP500

Propriété 3 (*Queues de distribution épaisses*) *L'hypothèse de normalité des rendements est généralement rejetée. Les queues des distributions empiriques des rendements sont généralement plus épaisses que celles d'une loi gaussienne. On parle alors de distribution leptokurtique.*

On rappelle que la Kurtosis d'une variable aléatoire X correspond à son moment centré d'ordre 4, i.e. $\mu_4 = E[(X - \mu)^4]$. La Kurtosis est une mesure de "l'épaisseur" des queues de distributions. En règle générale, on exprime cette mesure en contrôlant par une fonction puissance de la variance $V(X) = \sigma^2$. On définit ainsi une nouvelle mesure: le *degré d'excès de Kurtosis*.

$$\text{Degré d'excès de Kurtosis} = E \left[\left(\frac{X - E(X)}{\sigma_x} \right)^4 \right] - 3 \quad (2.5)$$

Cette dernière mesure est fondée par rapport à la distribution normale, qui est considérée comme une distribution à queue "plate", et qui possède un degré d'excès de Kurtosis normalisé à 0. Si la Kurtosis excède 3 (queues épaisses) la distribution est dite *leptokurtique*, si la Kurtosis est inférieure à 3, la distribution est dite *platikurtique*. Le moment empirique correspondant au degré d'excès de Kurtosis pour un échantillon de taille T s'écrit:

$$K_u = \frac{1}{T} \sum_{t=1}^T \left(\frac{x_t - \bar{x}}{\hat{\sigma}_x} \right)^4 \quad (2.6)$$

ou $\hat{\sigma}_x$ désigne un estimateur non biaisé de la variance. C'est ce moment qui est reporté sur la figure (2.4). Sous l'hypothèse nulle de normalité, on montre que:

$$\frac{K_u - 3}{\sqrt{\frac{24}{T}}} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

On vérifie ici que la kurtosis est largement et significativement supérieure à 3, impliquant l'existence d'une distribution leptokurtique des rendements de l'indice américain sur la période. Naturellement, le test de Jarque-Bera conduit ici à rejeter très largement l'hypothèse d'une distribution normale. Rappelons que ce test qui admet pour hypothèse nulle la normalité de la distribution est construit de la façon suivante:

$$s = \frac{T}{6} S_k + \frac{T}{24} (K_u - 3)^2 \xrightarrow[T \rightarrow \infty]{\mathcal{L}/H_0} \chi^2(2) \quad (2.7)$$

où S_k désigne la Skewness S_k qui sous l'hypothèse de symétrie est égale à 0.

Propriété 4 (*Clusters de Volatilité*) *On observe empiriquement que de fortes variations des rendements sont généralement suivies de fortes variations. On assiste ainsi à un regroupement des extrêmes en cluster ou paquets de volatilités.*

Figure 2.4: Histogramme des Rendements sur SP500

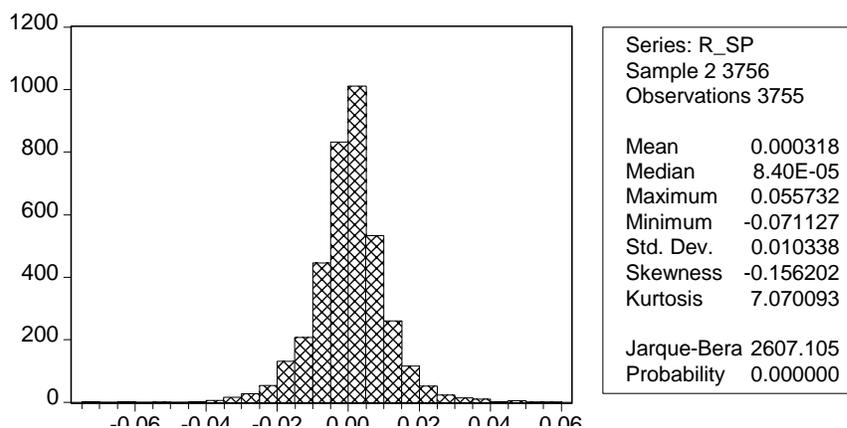
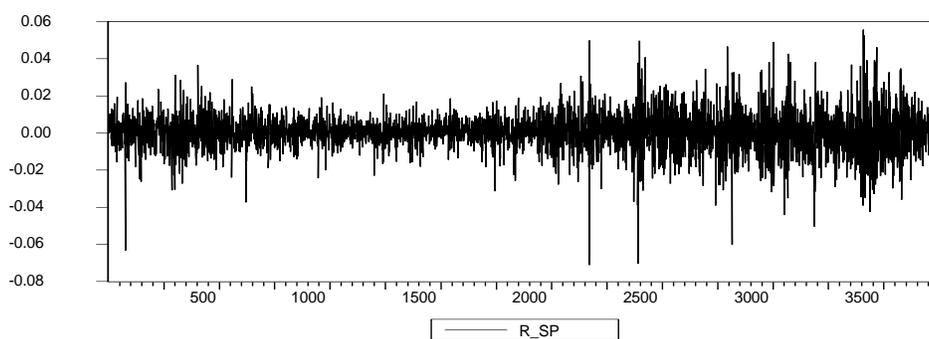


Figure 2.5: Illustration des Clusters de Volatilité sur les Rendements du SP500



On observe sans difficulté un tel phénomène sur les rendements du SP500 (figure 2.5). Naturellement, ce type de phénomène remet en cause l'hypothèse d'homoscédasticité généralement adoptée en économétrie linéaire.

Propriété 5 (*Queues épaisses conditionnelles*) *Même une fois corrigée de la volatilité clustering (par exemple avec des modèles ARCH), la distribution des résidus demeure leptokurtique même si la kurtosis est plus faible que dans le cas non conditionnelle.*

Propriété 6 (*Effet de levier*) *Il existe une asymétrie entre l'effet des valeurs passées négatives et l'effet des valeurs passées positives sur la volatilité des cours ou de*

rendements. Les baisses de cours tendent à engendrer une augmentation de la volatilité supérieure à celle induite par une hausse des cours de même ampleur.

Cette propriété n'est pas à confondre avec celle d'asymétrie de la distribution des cours ou des rendements (propriété 8). Il s'agit ici d'une asymétrie de la relation liant les valeurs passés des cours ou rendements à la volatilité de ces derniers.

Propriété 7 (Saisonnalité) *Les returns présentent de nombreux phénomènes de saisonnalité (effets week end, effet janvier etc..).*

Il existe une littérature très abondante qui tend à mettre en évidence de tels phénomènes. Toutefois, certaines saisonnalités peuvent être spécifiques à un échantillon, une période.. Il est, par contre, deux types de saisonnalité qui ont acquis droit de cité dans la littérature, pour avoir été discernés sur des échantillons, des périodes et au moyen de méthodologies présentant suffisamment de variété pour en étayer la robustesse. Il s'agit de "l'effet janvier" et de "l'effet week-end".

L'effet janvier apparaît très prononcé sur le marché américain comme l'ont montré notamment les études de Rozeff et Kinney (1976) et de Keim (1983). Ainsi, Rozeff et Kinney (1976) mettent en avant les résultats suivants sur le marché des actions US (tableau ??):

Tableau 2.1: Illustration de l'Effet Janvier.

Période	Return Moyen (% par mois)	
	Janvier	Autres Mois
1904-1928	1.30	0.44
1929-1940	6.63	-0.60
1940-1974	3.91	0.70
1904-1974	3.48	0.42

Sources: Rozeff et Kinney (1976) cité dans Cobbaut (1997)

Lorsque l'il croise cet effet janvier avec un "effet taille", Keim (1983) montre ainsi que l'effet janvier est d'autant plus fort que la capitalisation du titre est faible. On peut alors avancer une explication de type *tax loss selling hypothesis*. Les faibles capitalisations sont dans une large mesure le fait de sociétés dont le prix du titre a décliné de manière importante au cours de l'année, de sorte que les investisseurs ont intérêt à les vendre en fin d'année pour rendre leur moins value fiscalement déductible, quitte à les racheter quelques jours plus tard.

Autre saisonnalité prononcée : l'"effet" week end" étudié notamment par French (1980) ou Gibbon et Hess (1981). Il s'agit plus généralement d'un effet jour de la semaine comme l'atteste le tableau (2.2). Le phénomène le plus marquant est la différence entre les prix de clôture du vendredi et les cours d'ouverture du lundi. L'explication de ce phénomène est délicate. Une piste avancée consiste à postuler que lorsque les

entreprises ont de mauvaises nouvelles à annoncer, elles le font généralement le week end. Mais cette explication est a priori peu convaincante.

Tableau 2.2: Illustration de l'Effet Week-End.

		Lundi	Mardi	Mercredi	Jeudi	Vendredi
French (1980)	1953-1977	-0.17	0.02	0.10	0.04	0.09
Gibbons et Hess (1981)	1962-1978	-0.13	0.00	0.10	0.03	0.08

Sources: Cobbaut (1997). Return en pourcentage par mois, marché action US.

Propriété 8 (*Asymétrie perte/gain*) *La distribution des cours est généralement asymétrique : il y a plus de mouvements forts à la baisse qu'à la hausse.*

Rappelons qu'un test simple de l'hypothèse de symétrie consiste à tester la nullité du moment centré d'ordre trois de la distribution, i.e. la skewness.

$$Skewness = \mu_3 = E \left[(X - E(X))^3 \right] \quad (2.8)$$

Souvent on construit un coefficient de Skewness défini comme le rapport $S_K = \mu_3^2 / \sigma^3$. Sous l'hypothèse nulle de distribution normale et donc par conséquent de symétrie, on montre que:

$$\frac{(S_k)^{1/2}}{\sqrt{\frac{6}{T}}} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1) \quad (2.9)$$

On vérifie ainsi aisément sur la figure (2.4) que l'hypothèse nulle de symétrie de la distribution des rendements sur le SP500 est rejetée : le coefficient de skewness est significativement négative, dès lors la distribution est non symétrique, la probabilité d'obtenir des valeurs inférieures à la moyenne étant supérieure à celle d'obtenir des valeurs plus fortes que la moyenne. On retrouve la propriété d'asymétrie aux pertes et gains.

2.2. Les grandes classes de modèles non linéaires

Naturellement ces propriétés sont difficiles, voir impossibles, à reproduire à partir de modèle ARMA linéaires classiques. Ces modèles linéaires de séries temporelles n'étaient finalement fondés que sur des combinaisons linéaires de valeurs présentes et passés de chocs. En effet, le théorème central de l'analyse des séries temporelles qui est le théorème de Wold (1954), indique que tout processus faiblement stationnaire peut être réécrit sous la forme d'une moyenne mobile infinie de processus de type bruits blancs, c'est à dire sous la forme d'une combinaison linéaire d'une séquence de variable aléatoires non corrélées dans le temps. Par conséquent, l'hypothèse de processus ARMA stationnaire ne permet pas de prendre en compte d'une part les mécanismes d'asymétrie et d'autre part les ruptures de forte amplitude comme le montre la simulation suivante d'un processus $ARMA(2, 2)$. D'où la nécessité d'aller vers des modélisations non linéaires.

Insérer simulation ARMA : trajectoire + histo

Rappelons la définition générale d'un processus autorégressif et d'un processus linéaire autorégressif (Gourieroux, 1992).

Definition 2.3. *Un processus stochastique X_t est un processus autorégressif d'ordre K si et seulement si :*

$$E\left(X_t / \underline{X_{t-1}}\right) = E\left(X_t / X_{t-1}, X_{t-2}, \dots, X_{t-K}\right) \quad (2.10)$$

Un processus stochastique est un processus autorégressif linéaire d'ordre K si et seulement si :

$$EL\left(X_t / \underline{X_{t-1}}\right) = EL\left(X_t / X_{t-1}, X_{t-2}, \dots, X_{t-K}\right) \quad (2.11)$$

où $EL(.)$ désigne l'espérance linéaire.

L'espérance conditionnelle $E\left(X_t / \underline{X_{t-1}}\right)$ est la meilleure approximation au sens de l'erreur quadratique moyenne de Y_t par une fonction des valeurs passées X_{t-1}, X_{t-2}, \dots . Cette approximation est en général une fonction non linéaire de ces valeurs. La régression linéaire $EL\left(X_t / \underline{X_{t-1}}\right)$ est la meilleure approximation de X_t par une fonction linéaire affine de X_{t-1}, X_{t-2}, \dots . Au regard de cette définition, il existe une infinité de processus non linéaire susceptible de représenter les propriétés des séries financières.

Campbell, Lo et MacKinlay (1997) ont proposé le cadre suivant pour décrire un processus non-linéaire :

$$X_t = g(\varepsilon_{-1}, \varepsilon_{t-2}, \dots) + \varepsilon_t h(\varepsilon_{-1}, \varepsilon_{t-2}, \dots) \quad (2.12)$$

où la fonction $g(\cdot)$ correspond à la moyenne conditionnelle du processus X_t et où la fonction $h(\cdot)$ correspond à un coefficient de proportionnalité entre X_t et le choc ε_t . Cela permet de classer les processus non linéaires en deux parties :

1. Processus non-linéaires en moyenne pour lesquels $g(\cdot)$ est non-linéaire
2. Processus non-linéaires en variance pour lesquels $h(\cdot)$ est non-linéaire

Cette classification permet de regrouper la plupart des modèles non linéaires.

Une autre façon d'appréhender cette littérature sur les processus non linéaires consiste à opposer deux types d'approches. La première approche fondée **sur des extensions non linéaires de processus ARMA** qui permettent notamment d'appréhender les **mécanismes d'asymétrie et de seuil**. Pour spécifier ces phénomènes d'asymétrie et de seuils, les économètres ont développé toute une panoplie de spécifications :

- Modèles bilinéaires (Granger et Anderson, 1978)
- Modèles exponentiels autorégressifs (modèles EXPAR)
- Modèles à seuils (TAR, SETAR, STAR, MA asymétriques etc...) développés depuis les travaux pionniers de Tong (1978)
- Modèles MA non linéaires

La seconde voie a consisté à proposer une représentation autorégressive de la variance conditionnellement à son information passée permettant de tenir compte des phénomènes de **volatilité**. Dans ce domaine le papier de Engle de 1982, " *Autoregressive Conditional Heteroskedasticity with Estimates of the variance of UK inflation* ", *Econometrica* (1982) a ouvert la voie à la modélisation **ARCH** et à ses nombreux développements. C'est précisément sur cette voie que notre cours portera pour l'essentiel.

Mais avant cela, nous allons présenter un certain nombre de modèles non-linéaires appartenant à la première approche dont certains sont pourtant proches des modèles ARCH de par leurs implications.

2.2.1. Modèles bilinéaires (Granger et Andersen, 1978)

Le modèle bilinéaire, introduit par Granger et Andersen (1978), présente la particularité d'être à la fois linéaire en X_t et ε_t mais de ne pas l'être par rapport à ces deux variables prises conjointement. Un modèle bilinéaire d'ordre, noté $BL(p, q, P, Q)$, s'écrira ainsi sous la forme :

$$X_t = \mu + \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=0}^q \theta_i \varepsilon_{t-i} + \sum_{i=1}^P \sum_{j=1}^Q \lambda_{ij} X_{t-i} \varepsilon_{t-j} \quad (2.13)$$

avec $\theta_0 = 1$ et $(\phi_p, \theta_q, \lambda_{iQ}, \lambda_{Pj}) \in \mathbb{R}^{*4}, \forall (i, j)$ et où ε_t désigne un bruit blanc, éventuellement gaussien.

Certains des processus bilinéaires ont des propriétés proches de celles des modèles ARCH que nous étudierons dans ce chapitre. Considérons un exemple particulier de processus $BL(0, 0, 2, 1)$ de type:

$$X_t = \varepsilon_t + \lambda X_{t-2} \varepsilon_{t-1} \quad (2.14)$$

où $\lambda \in \mathbb{R}$ et où ε_t est *i.i.d.* $(0, \sigma_\varepsilon^2)$. Ce processus est de moyenne nulle car:

$$\begin{aligned} E(X_t) &= E(\varepsilon_t) + \lambda E(X_{t-2} \varepsilon_{t-1}) \\ &= E(\varepsilon_t) + \lambda E(X_{t-2}) E(\varepsilon_{t-1}) = 0 \end{aligned}$$

Sa fonction d'autocovariance $\gamma(h) = E(X_t X_{t-h})$ est définie de la façon suivante:

$$\begin{aligned} \gamma(h) &= E(\varepsilon_t + \lambda X_{t-2} \varepsilon_{t-1})(\varepsilon_{t-h} + \lambda X_{t-2-h} \varepsilon_{t-1-h}) \\ &= E(\varepsilon_t \varepsilon_{t-h}) + \lambda^2 E(X_{t-2} \varepsilon_{t-1} X_{t-2-h} \varepsilon_{t-1-h}) \\ &\quad + \lambda E(X_{t-2} \varepsilon_{t-1} \varepsilon_{t-h}) + \lambda E(\varepsilon_t X_{t-2-h} \varepsilon_{t-1-h}) \end{aligned}$$

Pour $h > 0$, il n'apparaît aucun terme en ε_{t-h}^2 et puisque l'opérateur espérance est linéaire, la fonction $\gamma(h)$ est par conséquent nulle. En revanche, pour $h = 0$, on a :

$$\begin{aligned} \gamma(0) &= E(\varepsilon_t^2) + \lambda^2 E(X_{t-2}^2) E(\varepsilon_{t-1}^2) \\ &= \sigma_\varepsilon^2 + \lambda^2 E(X_{t-2}^2) \sigma_\varepsilon^2 \end{aligned}$$

Ainsi la fonction génératrice d'autocovariance s'écrit :

$$E(X_t X_{t-h}) = \begin{cases} \sigma_\varepsilon^2 + \lambda^2 E(X_{t-2}^2) \sigma_\varepsilon^2 & \text{si } h = 0 \\ 0 & \text{si } h \geq 1 \end{cases} \quad (2.15)$$

Examinons la variance et la variance conditionnelle du processus X_t défini par ce processus bilinéaire $BL(0, 0, 2, 1)$. La variance du processus X_t sous la condition d'érgodicité $\lambda^2 \sigma_\varepsilon^2 < 1$ est obtenue par la solution de l'équation de récurrence :

$$\sigma_{X,t}^2 - (\lambda^2 \sigma_\varepsilon^2) \sigma_{X,t-2}^2 - \sigma_\varepsilon^2 = 0$$

On montre alors que la variance marginale s'écrit $Var(X_t) = \sigma_\varepsilon^2 / (1 - \lambda^2 \sigma_\varepsilon^2)$. Parallèlement, la variance conditionnelle du processus X_t se dérive directement à partir de l'équation (2.14).

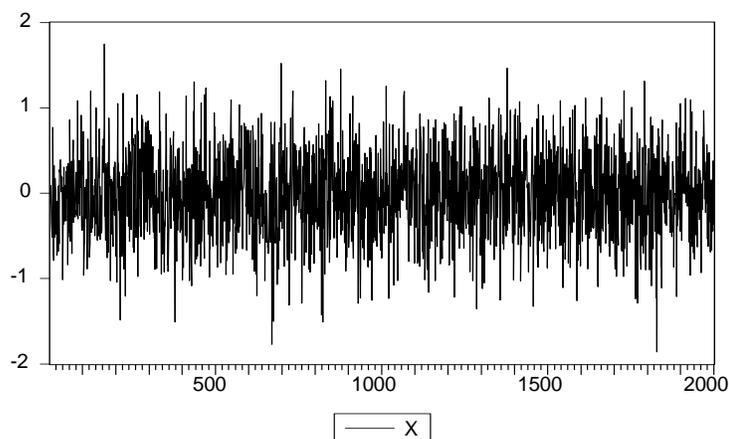
$$V(X_t / X_{t-2}) = \sigma_\varepsilon^2 [1 + \lambda^2 X_{t-2}^2] \quad (2.16)$$

La variance conditionnelle du processus X_t dépend des valeurs passées de ce processus. On retrouve un effet type ARCH. Ceci illustre le fait que plusieurs modélisations

non linéaires peuvent être envisagées si l'on souhaite modéliser la dynamique dans la volatilité conditionnelle.

On vérifie sur le graphique (2.6) que le modèle bilinéaire $BL(0,0,2,1)$ avec $\lambda = 0.2$ est capable de générer des cluster de volatilité comme ceux observés sur données financières.

Figure 2.6: Simulation d'un Processus $BL(0,0,2,1)$



2.2.2. Modèles auto-régressifs exponentiels (modèles EXPAR)

Ces modèles constituent ont une structure autorégressive de type AR, mais permettent de prendre en compte des phénomènes de cluster de volatilité de la série. Un modèle EXPonentiel AutoRegresif (EXPAR) est défini par la relation :

$$X_t = \mu + \sum_{i=1}^p [\alpha_i + \beta_i \exp(-\gamma X_{t-1}^2)] X_{t-i} + \varepsilon_t \quad (2.17)$$

On reconnaît ici une structure multiplicative sensiblement proche des modèles bilinéaires. Une simulation de type de processus est reproduit sur la figure ().

2.2.3. Modèles autorégressifs à seuil (modèles TAR)

Les modèles auto-régressifs à seuils constituent l'une des spécifications possibles de la grande famille des modèles non-linéaires, appelés modèles à régime. L'idée consiste à postuler l'existence de plusieurs dynamiques pour une même série (*plusieurs régimes*) et à spécifier un mécanisme de transition d'un régime à l'autre. Deux grands types de mécanismes de transition existent :

- Des mécanismes de transition stochastiques et exogènes régis par des processus de type chaîne de Markov : on parle alors de *Markov Switching Models* (Hamilton, 1989)
- Des mécanismes de transition endogènes où la fonction de transition dépend de la variable dépendante et d'un seuil : on parle alors de *Modèles à Seuils* ou *Threshold AutoRegressive Models (TAR)*

Les modèles à seuils ont été introduit par Tong (1978). Il existe toute une classe de modèle suivant la définition retenue de la fonction de transition : TAR, MTAR, STAR, ESTAR, LSTAR, MSTAR etc.. Le modèle le plus simple est le modèle SETAR (*Self Exciting Threshold AutoRegressive*) introduit par Tong, mais popularisé par Hansen (1996). Considérons le cas le plus simple d'un modèle à deux régimes :

$$X_t = \Phi_1(L) X_t \mathbb{I}_{X_{t-d} > \gamma} + \Phi_2(L) X_t \mathbb{I}_{X_{t-d} \leq \gamma} + \varepsilon_t \quad (2.18)$$

où $\Phi_j(L)$, $j = 1, 2$ désignent deux polynômes retard d'ordre fini et où ε_t est *i.i.d.* $(0, \sigma_\varepsilon^2)$. La fonction \mathbb{I}_z désigne l'indicatrice telle que :

$$\mathbb{I}_{X_{t-d} > \gamma} = \begin{cases} 1 & \text{si } X_{t-d} > \gamma \\ 0 & \text{sinon} \end{cases} \quad (2.19)$$

Le paramètre $d \in \mathbb{N}$, appelé délai, est nécessairement supérieure à l'unité pour éviter des problèmes de simultanéité. Le paramètre $\gamma \in \mathbb{R}$ est appelé paramètre de seuil. Ce type de modèle permet très facilement de modéliser des phénomènes tels que l'asymétrie : pour un même choc, les mécanismes de propagation diffèrent suivant les valeurs passées de la variable dépendante X_{t-d} . Ce type de processus permet aussi d'obtenir des distributions leptokurtiques de la variable dépendante.

Insérer simulation d'un processus à seuil

Ainsi, cette présentation de quelques grandes classes de modèles non linéaires montre qu'il existe un grand nombre de voies de recherche possibles lorsque l'on cherche à reproduire au mieux les principales propriétés des séries financières. L'approche ARCH / GARCH est une approche en particulier qui doit donc être repensé dans le cadre plus vaste des modèles non linéaires comme l'indique la citation de Berra et Higgins (1993) du début de cette section.

2.3. L'approche ARCH / GARCH et la modélisation de l'incertitude

L'approche ARCH / GARCH, comme nous l'avons dit précédemment, a été proposée pour prendre en compte des variances conditionnelles dépendant du temps. Le principe général consiste donc à remettre en cause la propriété d'**homoscédasticité** que l'on retient généralement dans le cadre du modèle linéaire.

”Commençons par un petit exemple introductif. Dans l’analyse traditionnelle de la prévision (cf. Box et Jenkins), la construction des valeurs futures prévues est fondée sur la moyenne conditionnelle de la série utilisée. Ainsi, la prévision de la valeur d’une série temporelle y_t à la date $t+1$ compte tenu de l’information disponible à la date t est donnée par $E(y_{t+1}/y_t, y_{t-1}, \dots)$. Supposons que y_t suivent un processus $AR(1)$ stationnaire $y_t = \theta y_{t-1} + \varepsilon_t$, avec ε_t *i.i.d* $(0, \sigma_\varepsilon^2)$, alors la moyenne conditionnelle de y_{t+1} est θy_t tandis que son espérance inconditionnelle est nulle. Comme le fait remarquer Engle (1982), l’amélioration des prévisions issues de modèles de séries temporelles provient clairement de l’exploitation de l’information contenue dans **l’espérance conditionnelle** du processus. L’idée serait tenir compte des autres moments conditionnels de ce processus. Or, la variance conditionnelle du processus $AR(1)$ est égale à σ_ε^2 , tandis que sa variance inconditionnelle est égale à $\sigma_\varepsilon^2 / (1 - \theta)$. Ainsi les variances de prévision conditionnelles et inconditionnelles sont constantes quelle que soit la date de la prévision. Avec de tels modèles on est donc incapables de mesurer d’éventuels changements dans les variances des erreurs de prévision même si l’on souhaite que celles-ci soient affectées par l’évolution passée. Or par exemple en théorie financière *l’étude du CAPM implique que l’espérance conditionnelle du rendement d’un actif par rapport à un actif non risqué soit proportionnelle à la covariance conditionnelle entre ce rendement et celui d’un portefeuille comprenant l’ensemble des titres du marché*. Dès lors, une telle théorie **impose une paramétrisation des variances conditionnelles** : la variance conditionnelle doit pouvoir évoluer avec le temps.

Si l’on dit à l’économètre classique qu’il doit estimer les paramètres d’un modèle dont la variance des perturbations évolue avec le temps, il évoquera un problème d’**hétéroscédasticité**, qui signifie tout simplement que la matrice de variance covariance des erreurs n’est pas définie à un scalaire près par la matrice identité. Autrement dit, les termes de la diagonale principale de la matrice de variance covariance (c’est à dire les variances) ne sont pas tous identiques pour les différentes perturbations intervenant à des dates différentes. Il convient alors de traiter ce problème pour obtenir des estimateurs efficaces : généralement, la solution consiste à introduire une variable exogène x_t qui permet de prévoir l’évolution de la variance. Si par exemple $y_t = \varepsilon_t \cdot x_{t-1}$, alors la variance conditionnelle de y_t est égale à $\sigma_\varepsilon^2 \sigma_{x,t-1}^2$. Les intervalles de prévision sur y_t dépendront alors de l’évolution de la variable exogène x_t . Toutefois, cette première approche est peu satisfaisante puisqu’elle nécessite de spécifier a priori une cause à l’évolution de la variance.

Le principe général proposé par **Engle (1982)** consiste à supposer que la variance dépend de l’ensemble informationnel dont on dispose. Il propose une spécification $ARCH(q)$ où le carré des perturbations suit un processus autorégressif d’ordre q . **Les modèles ARCH sont donc des modèles autorégressifs conditionnellement hétéroscédastiques**. Engle (1982) a donc proposer ces processus pour palier aux

insuffisances de la classe des représentations ARMA, notamment en ce qui concerne les séries financières qui présentent une volatilité (ou variabilité instantanée mesurée par la variance conditionnelle) fonction du temps et par des ajustements asymétriques. Ainsi, **les modèles ARCH sont basés sur une paramétrisation endogène de la variance conditionnelle**. La famille des modèles ARCH peut se décomposer en deux sous-ensembles : les modèles ARCH linéaires et les modèles ARCH non linéaires. Les premiers reposent sur une spécification quadratique de la variance conditionnelle des perturbations : modèles $ARCH(q)$, $GARCH(p, q)$ et $IGARCH(p, q)$. Les modèles ARCH non linéaires sont caractérisés par des spécifications asymétriques des perturbations. Ce sont les modèles $EGARCH(p, q)$, $TARCH(q)$ et $TGARCH(p, q)$." (Bresson et Pirotte, *Séries temporelles*)

3. Modèles ARCH / GARCH linéaires

3.1. Modèles ARCH(q)

Commençons par présenter le modèle ARCH(1) introduit Engle (1982). On considère un processus X_t tel que :

$$X_t = z_t \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} \quad (3.1)$$

où z_t est un bruit blanc faible tel que $E(z_t^2) = \sigma_z^2$. On note généralement $h_t = \alpha_0 + \alpha_1 X_{t-1}^2$ et donc par conséquent un processus ARCH(1) s'écrit sous la forme suivante.

Definition 3.1. *Un processus X_t satisfait une représentation ARCH(1) si*

$$X_t = z_t \sqrt{h_t} \quad (3.2)$$

avec $h_t = \alpha_0 + \alpha_1 X_{t-1}^2$ et où z_t désigne un bruit blanc faible tel que $E(z_t) = 0$ et $E(z_t^2) = \sigma_z^2$.

De façon générale, z_t désigne un ensemble de variables aléatoires indépendantes, identiquement distribuées, centrées, réduites. La composante h_t désigne une variable qui, conditionnellement à l'ensemble d'information des valeurs passées de X_t , i.e. à $\underline{X}_{t-1} = \{X_{t-1}, X_{t-2}, \dots, X_{t-j}, \dots\}$, est déterministe et positive. Dans ce système, le processus X_t est caractérisé par des autocorrélations nulles et une variance conditionnelle variable dans le temps en fonction de l'ampleur des de l'innovation passée.

On peut établir des résultats intéressants en considérant le processus autorégressif sur X_t^2 . Pour simplifier, on se limite au cas du ARCH(1). Dans ces conditions :

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 \Leftrightarrow X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + (X_t^2 - h_t) \quad (3.3)$$

soit encore :

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \varepsilon_t \quad (3.4)$$

ou $\varepsilon_t = (X_t^2 - h_t)$ vérifiant $E(\varepsilon_t | X_{t-1}) = 0$ est un processus d'innovation pour X_t^2 . Ainsi, cette écriture précédente correspond çà celle d'un AR(1) sur le carré X_t^2 . On sait que ce processus X_t^2 est stationnaire au second ordre si et seulement si $\alpha_1 < 1$. On peut déduire de ces différentes écritures, un certain nombre de propriétés qui pourront être étendues au cas des processus ARCH(q).

On vérifie tout d'abord que le processus X_t vérifie la définition d'une différence de martingale homoscedastique et d'un bruit blanc faible. Rappelons au préalable ces deux définitions :

Definition 3.2. *Un processus X_t est un bruit blanc faible si il s'agit d'une suite de variables de moyenne nulle, homoscedastiques et non corrélés. Si l'on note $EL(\cdot)$ l'espérance linéaire :*

$$EL(X_t | X_{t-1}) = 0 \quad V(X_t) = \sigma_x^2 \quad \forall t \quad (3.5)$$

Un processus X_t est une différence de martingale homoscedastique si et seulement si :

$$E(X_t | X_{t-1}) = 0 \quad V(X_t) = \sigma_x^2 \quad \forall t \quad (3.6)$$

Propriété 1 *Le processus X_t ARCH(1) défini par l'équation (3.2) est une différence de martingale homoscedastique :*

$$E(X_t | X_{t-1}) = 0 \quad V(X_t) = \frac{\alpha_0}{1 - \alpha_1} \quad \forall t \quad (3.7)$$

Cette propriété signifie que le processus ARCH X_t qui peut s'apparenter à un processus de bruit blanc (faible), ce qui explique notamment que l'on spécifiera des erreurs de modèles sous la forme ARCH. On retrouve alors toutes les propriétés de modèles établies sous la propriété de bruit blanc des erreurs. Mais cette propriété signifie en outre que le processus ARCH X_t est non conditionnellement homoscedastique.

Preuve : nous ne démontrerons que la première partie de la propriété 1, le reste étant démontré par la suite.

$$E(X_t | X_{t-1}) = E(z_t \sqrt{h_t} | X_{t-1}) = E(z_t | X_{t-1}) \sqrt{h_t} = 0$$

si le processus z_t est lui même un bruit blanc faible. On peut en outre montrer par itération que le processus X_t est centré, i.e. $E(X_t) = 0$.

Par développement, on peut montrer que le processus ARCH X_t est orthogonal à tout passé :

$$E(X_t | X_{t-h}) = 0 \quad \forall h \geq 1 \quad (3.8)$$

Pour démontrer cela utilisons la propriété des espérances itérées. Si l'on considère deux ensembles d'information Ω_1 et Ω_2 , tel que $\Omega_1 \subseteq \Omega_2$, alors quelle que soit la variable Z , on a l'égalité $E(Z|\Omega_1) = E[E(Z|\Omega_2)|\Omega_1]$. En effet, on montre que

$$E\left(X_t|\underline{X}_{t-h}\right) = E\left[E\left(X_t|\underline{X}_{t-1}\right)|\underline{X}_{t-h}\right] = E\left[0|\underline{X}_{t-h}\right] = 0$$

Propriété 2 La variance conditionnelle du processus X_t ARCH(1) défini par l'équation (3.2) est non constante dans le temps et vérifie :

$$V\left(X_t|\underline{X}_{t-h}\right) = \alpha_0 \left(\frac{1 - \alpha_1^h}{1 - \alpha_1}\right) + \alpha_1^h X_{t-h}^2 \quad \forall t \quad (3.9)$$

C'est la propriété centrale des processus ARCH : le processus X_t possède les propriétés d'un bruit blanc homoscédastique (propriété 1), mais sa variance conditionnelle dépend du temps. On a l'idée que la liaison temporelle passe par l'intermédiaire de l'équation auto-régressive définie sur le carré du processus, i.e. $X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \varepsilon_t$.

Preuve : On sait que $E\left(X_t|\underline{X}_{t-h}\right) = 0$, dès lors $V\left(X_t|\underline{X}_{t-h}\right) = E\left(X_t^2|\underline{X}_{t-h}\right)$. Considérons le processus X_t^2 défini par la relation $X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \varepsilon_t$ où ε_t est un bruit blanc faible. Par itération successive, on a :

$$\begin{aligned} X_t^2 &= \alpha_0 \left(1 + \alpha_1 + \alpha_1^2 + \dots + \alpha_1^h\right) + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_1^2 \varepsilon_{t-2} + \dots \\ &\quad + \alpha_1^{h-1} \varepsilon_{t-h+1} + \alpha_1^h X_{t-h}^2 \end{aligned} \quad (3.10)$$

En considérant l'espérance conditionnelle de chacun de ces membres, il vient :

$$E\left(X_t^2|\underline{X}_{t-h}\right) = \alpha_0 \left(\frac{1 - \alpha_1^h}{1 - \alpha_1}\right) + \sum_{j=0}^{h-1} \alpha_1^j E\left(\varepsilon_{t-j}|\underline{X}_{t-h}\right) + \alpha_1^h E\left(X_{t-h}^2|\underline{X}_{t-h}\right)$$

Puisque par définition du bruit blanc ε_t , on a $E\left(\varepsilon_{t-j}|\underline{X}_{t-h}\right) = 0$, $\forall j = 0, \dots, h-1$ et par définition $E\left(X_{t-h}^2|\underline{X}_{t-h}\right) = X_{t-h}^2$, on obtient ainsi la formule de la proposition 2.

$$V\left(X_t|\underline{X}_{t-h}\right) = \alpha_0 \left(\frac{1 - \alpha_1^h}{1 - \alpha_1}\right) + \alpha_1^h X_{t-h}^2 \quad \forall t \quad (3.11)$$

Lorsque h tend vers l'infini, ces variances conditionnelles convergent vers la variance non conditionnelle, et l'on retrouve alors la formule de la propriété 1 :

$$V(X_t) = \lim_{h \rightarrow \infty} V\left(X_t|\underline{X}_{t-h}\right) = \lim_{h \rightarrow \infty} \left[\alpha_0 \left(\frac{1 - \alpha_1^h}{1 - \alpha_1}\right) + \alpha_1^h X_{t-h}^2 \right] = \frac{\alpha_0}{1 - \alpha_1} \quad (3.12)$$

Propriété 3 Les auto-covariances conditionnelles du processus X_t ARCH(1) défini par l'équation (3.2) sont nulles :

$$\text{cov} \left(X_t, X_{t+k} | \underline{X_{t-h}} \right) = 0 \quad \forall h \geq 1, \forall k \geq 1 \quad (3.13)$$

Un processus est donc un processus qui conditionnellement à $\underline{X_{t-h}}$ est un processus sans mémoire.

Preuve : Considérons la covariance conditionnelle entre X_t et X_{t+k} :

$$\begin{aligned} \text{cov} \left(X_t, X_{t+k} | \underline{X_{t-h}} \right) &= E \left\{ \left[X_t X_{t+k} - E \left(X_t | \underline{X_{t-h}} \right) E \left(X_{t+k} | \underline{X_{t-h}} \right) \right] | \underline{X_{t-h}} \right\} \\ &= E \left(X_t X_{t+k} | \underline{X_{t-h}} \right) \\ &= E \left[E \left(X_t X_{t+k} | \underline{X_{t+k-1}} \right) | \underline{X_{t-h}} \right] \\ &= E \left[X_t E \left(X_{t+k} | \underline{X_{t+k-1}} \right) | \underline{X_{t-h}} \right] \quad \text{car } \varepsilon_t \text{ est connu en } t+k-1 \\ &= E \left[X_t \times 0 | \underline{X_{t-h}} \right] \end{aligned}$$

Propriété 4 Les conditions suffisantes qui garantissent la positivité du processus X_t^2 sont $\alpha_1 > 0$ et $\alpha_0 + \varepsilon_t \geq 0$ pour toute valeur admissible de ε_t . Ceci implique notamment des contraintes sur le support de la loi de ε_t . De plus, la variance marginale du processus X_t existe si et seulement si $\alpha_0 > 0$ et $0 < \alpha_1 \leq 1$.

En effet, il convient de vérifier notamment que $V(X_t^2)$ et $V(X_t)$ sont définies de façon positive. Sous les conditions $\alpha_0 > 0$ et $0 < \alpha_1 \leq 1$, la variance de X_t existe et est constante dans le temps : le processus X_t est stationnaire au second ordre.

On peut en outre établir les moments conditionnels et non conditionnels d'ordre 4 du processus X_t .

Propriété 5 Le moment conditionnel centré d'ordre 4 du processus X_t vérifie :

$$E \left(X_t^4 | \underline{X_{t-h}} \right) = 3 \left(\alpha_0 + \alpha_1 X_{t-1}^2 \right)^2 \quad (3.14)$$

Sous l'hypothèse $3\alpha_1^2 < 1$, le moment non conditionnel centré d'ordre 4 du processus X_t est égal à :

$$E(X_t^4) = 3 \left[\alpha_0^2 + \frac{2\alpha_1\alpha_0^2}{1-\alpha_1} + \alpha_1^2 E(X_{t-1}^4) \right] = \frac{3\alpha_0^2(1+\alpha_1)}{(1-3\alpha_1^2)(1-\alpha_1)} \quad (3.15)$$

La kurtosis non conditionnelle associée au processus ARCH(1) est égale à :

$$\text{Kurtosis} = \frac{E(X_t^4)}{E(X_t^2)^2} = 3 \left(\frac{1-\alpha_1^2}{1-3\alpha_1^2} \right) > 3 \quad (3.16)$$

Preuve : cf. Berra et Higgins (1993).

Sous l'hypothèse de positivité du paramètre α_1 , la kurtosis non conditionnelle est toujours supérieure à celle de la loi normale : elle traduit l'aspect leptokurtique de la distribution du processus X_t . C'est donc la deuxième raison, avec la variance conditionnelle dépendante du temps, pour laquelle les processus ARCH sont très utilisés pour représenter les séries financières ou les résidus de modèles linéaires définis sur séries financières.

Toutes ces propriétés peuvent être généralisées au cas d'un processus ARCH(q) :

Definition 3.3. *Un processus X_t satisfait une représentation ARCH(q) si*

$$X_t = z_t \sqrt{h_t} \quad (3.17)$$

avec

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2$$

et où z_t désigne un bruit blanc faible tel que $E(z_t) = 0$ et $E(z_t^2) = \sigma_z^2$.

Pour ce type de processus, on retrouve les deux propriétés essentielles vues précédemment, à savoir la propriété de différence de martingale (ou bruit blanc faible) $E(X_t | X_{t-1}) = 0$ et la propriété de variance conditionnelle variable dans le temps puisque :

$$V(X_t | X_{t-1}) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 \quad (3.18)$$

3.2. Modèle avec erreurs ARCH(q)

On considère dorénavant non plus un processus ARCH pour modéliser directement la série financière, mais le résidu d'un modèle linéaire. Prenons l'exemple d'un modèle linéaire auto-régressif avec résidus de type ARCH(q).

Definition 3.4. *On considère un modèle linéaire auto-régressif de la forme*

$$Y_t = E(Y_t | Y_{t-1}) + \varepsilon_t \quad (3.19)$$

où ε_t est un bruit blanc faible, tel que $E(\varepsilon_t) = 0$ et $E(\varepsilon_t \varepsilon_s) = 0$ si $s \neq t$, satisfaisant la condition de différence de martingale $E(\varepsilon_t | \varepsilon_{t-1}) = 0$. On suppose que ce résidu admet une représentation de type ARCH(q) :

$$\varepsilon_t = z_t \sqrt{h_t} \quad \text{avec} \quad h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad (3.20)$$

où z_t est un bruit blanc faible.

On a donc un modèle qui décrit à la fois l'évolution de l'espérance conditionnelle et de la variance conditionnelle du processus Y_t dans le temps. Envisageons le cas le plus simple d'un processus de type AR(1) avec erreur ARCH(1):

$$Y_t = \mu + \rho Y_{t-1} + \varepsilon_t \quad (3.21)$$

$$\varepsilon_t = z_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \quad (3.22)$$

avec $|\rho| < 1$. Dans ce cas, les résidu satisfont les 4 principales propriétés étudiées précédemment :

- Propriété de différence de martingale : $E(\varepsilon_t | \varepsilon_{t-1}) = 0$ et de façon générale

$$E(\varepsilon_t | \varepsilon_{t-h}) = 0, \forall h \geq 1$$

- Variance conditionnelle dépendante du temps puisque :

$$V(\varepsilon_t | \varepsilon_{t-h}) = \alpha_0 \left(\frac{1 - \alpha_1^h}{1 - \alpha_1} \right) + \alpha_1^h \varepsilon_{t-h}^2 \quad V(\varepsilon_t) = \frac{\alpha_0}{1 - \alpha_1} \quad (3.23)$$

- Les auto-covariances conditionnelles sont nulles $cov(\varepsilon_t, \varepsilon_{t+k} | \varepsilon_{t-h}) = 0, \forall h \geq 1, \forall k \geq 1$
- Sous l'hypothèse $\alpha_1^2 < 1/3$, la distribution des résidus est leptokurtique puisque :

$$Kurtosis = 3 \left(\frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} \right) > 3 \quad (3.24)$$

On peut en outre en déduire un certain nombre de conclusions quant au processus Y_t lui même. On peut montrer tout d'abord que l'espérance conditionnelle de Y_t vérifie :

$$E(Y_t | Y_{t-h}) = \mu + \rho E(Y_{t-1} | Y_{t-h}) = \mu \left(\frac{1 - \rho^h}{1 - \rho} \right) + \rho^h Y_{t-h}$$

De la même façon, on montre que la variance conditionnelle de Y_t dépend du temps. En effet, on peut montrer qu'elle dépend du processus ε_{t-h}^2 de la façon suivante.

Propriété 1 La variance conditionnelle du processus AR(1) avec erreur ARCH(1), Y_t , s'écrit :

$$V(Y_t | Y_{t-h}) = \left(\frac{\mu}{1 - \alpha_1} \right) \left[\left(\frac{1 - \rho^{2h}}{1 - \rho^2} \right) - \alpha_1 \left(\frac{\alpha_1^h - \rho^{2h}}{\alpha_1 - \rho^2} \right) \right] + \alpha_1 \left(\frac{\alpha_1^h - \rho^{2h}}{\alpha_1 - \rho^2} \right) \varepsilon_{t-h}^2 \quad (3.25)$$

Ainsi la variance d'une erreur de prévision à l'horizon 1, s'écrit :

$$V(Y_t | Y_{t-1}) = \mu + \alpha_1 \varepsilon_{t-1}^2 \quad (3.26)$$

En conclusion, si l'on désire prévoir le processus Y_t dans le cas d'erreur ARCH(1), l'erreur de prévision à une horizon d'une période admet une variance $V\left(Y_t|Y_{t-1}\right)$ qui varie dans le temps en fonction de la valeur de ε_{t-1}^2 . Dès lors, les intervalles de confiance sur cet prévision ne sont plus constants dans le temps.

Remarque *La variance d'une erreur de prévision sur un processus avec erreur ARCH dépend du temps.*

$$V\left(Y_t|Y_{t-h}\right) = g\left(\varepsilon_{t-h}\right) \quad (3.27)$$

L'amplitude des intervalles de confiance associés à cette prévision n'est donc pas constante dans le temps.

3.3. Modèles GARCH(p, q)

Par la suite de cet exposé, on continue de considérer un modèle linéaire autorégressif exprimé sous la forme suivante :

$$Y_t = E\left(Y_t|Y_{t-1}\right) + \varepsilon_t \quad (3.28)$$

où ε_t est un bruit blanc faible, tel que $E(\varepsilon_t) = 0$ et $E(\varepsilon_t\varepsilon_s) = 0$ si $s \neq t$, satisfaisant la condition de différence de martingale

$$E\left(\varepsilon_t|\varepsilon_{t-1}\right) = 0 \quad (3.29)$$

On suppose toujours que le processus ε_t peut s'écrire sous la forme :

$$\varepsilon_t = z_t\sqrt{h_t} \quad (3.30)$$

où z_t est un bruit blanc faible.

On cherche à modéliser la volatilité conditionnelle du processus de bruit ε_t . Pour tenir compte de la dynamique observée sur ε_t^2 , on peut être amené à imposer une valeur élevée du paramètre q dans la modélisation ARCH(q) ce qui peut poser des problèmes d'estimation. Il s'agit d'une difficulté semblable à celle que l'on rencontre dans les modélisation de l'espérance conditionnelle : si le théorème de Wold assure que toute série stationnaire possède une représentation de type MA, il est possible que pour une série donnée, l'ordre de cet MA soit particulièrement élevé, voire infini. Dans ce cas, Box et Jenkins proposent de regagner en parcimonie en utilisant une représentation de type AR(p) ou ARMA(p, q). Pour la variance conditionnelle, Bollerslev (1986) définit ainsi le processus GARCH(p, q) en substituant à l'équation (3.20) l'expression :

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (3.31)$$

avec les conditions $\alpha_0 > 0$, $\alpha_i \geq 0, i = 1, \dots, q$ et $\beta_i \geq 0, i = 1, \dots, p$ suffisantes pour garantir la positivité de h_t .

Definition 3.5. Un processus ε_t satisfait une représentation GARCH(p, q) si

$$\varepsilon_t = z_t \sqrt{h_t} \quad (3.32)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (3.33)$$

où z_t est un bruit blanc faible et où $\alpha_0 > 0$, $\alpha_i \geq 0, i = 1, \dots, q$ et $\beta_i \geq 0, i = 1, \dots, p$.

Ainsi, l'erreur du processus Y_t , définie par le processus GARCH(p, q) ε_t admet pour moments conditionnels :

$$E(\varepsilon_t | \underline{\varepsilon}_{t-1}) = 0 \quad (3.34)$$

$$V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (3.35)$$

Tout comme pour le modèle ARCH, on peut par inversion exprimer le processus ε_t^2 sous la forme d'un processus ARMA défini dans une innovation

$$\mu_t = \varepsilon_t^2 - h_t \quad (3.36)$$

En introduisant cette notation dans l'équation (3.33), il vient :

$$\varepsilon_t^2 - \mu_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i (\varepsilon_{t-i}^2 - \mu_{t-i})$$

D'où l'on tire que :

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + \mu_t - \sum_{i=1}^p \beta_i \mu_{t-i}$$

avec la convention $\alpha_i = 0$ si $i > q$ et $\beta_i = 0$ si $i > p$.

Remarque Le processus ε_t^2 d'une représentation GARCH(p, q) peut être représenté sous la forme d'un processus ARMA[$\max(p, q), p$] défini dans une innovation $\mu_t = \varepsilon_t^2 - V(\varepsilon_t | \underline{\varepsilon}_{t-1})$, tel que :

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + \mu_t - \sum_{i=1}^p \beta_i \mu_{t-i} \quad (3.37)$$

avec la convention $\alpha_i = 0$ si $i > q$ et $\beta_i = 0$ si $i > p$.

Attention, le degré p ici apparaît comme le degré moyenne mobile de la représentation ARMA dans ε_t^2 . A partir de cette représentation, on peut calculer de façon assez simple les moments et moments conditionnels du processus d'erreur ε_t mais aussi du processus Y_t .

Exemple 1 : Considérons le cas d'un processus GARCH(1, 1) :

$$\varepsilon_t = z_t \sqrt{h_t} \quad (3.38)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (3.39)$$

qui peut être représenté par le modèle suivant :

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + \mu_t - \beta_1 \mu_{t-1} \quad (3.40)$$

où $\mu_t = \varepsilon_t^2 - V(\varepsilon_t/\varepsilon_{t-1}) = \varepsilon_t^2 - h_t$ est un processus d'innovation pour ε_t^2 . Sous la condition de stationnarité du second ordre $\alpha_1 + \beta_1 < 1$, la variance non conditionnelle du processus ε_t est définie et constante dans le temps. Sachant que $V(\varepsilon_t) = E(\varepsilon_t^2)$ il suffit à partir de la forme ARMA(1,1) sur ε_t^2 de définir l'espérance du processus :

$$V(\varepsilon_t) = E(\varepsilon_t^2) = \alpha_0 \Phi(1)^{-1} = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \quad (3.41)$$

où $\Phi(L) = 1 - (\alpha_1 + \beta_1)L$ désigne le polynôme auto-régressif associé à la représentation ARMA(1,1) sur ε_t^2 .

Enfin, on peut montrer que pour un processus GARCH la kurtosis est directement liée à l'hétéroscédasticité conditionnelle. Considérons le cas de la kurtosis associée à la loi non conditionnelle dans un processus GARCH conditionnellement gaussien :

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ N.i.d. } (0, 1) \quad (3.42)$$

Dans ce cas, les moments conditionnels d'ordre 2 et 4 du processus ε_t sont liés :

$$E(\varepsilon_t^4 | \varepsilon_{t-1}) = 3 \left[E(\varepsilon_t^2 | \varepsilon_{t-1}) \right]^2 \quad (3.43)$$

En effet, on rappelle que si une variable centrée y suit une loi gaussienne, $E(y^4) = 3\text{Var}(y)^2 = 3E(y^2)^2$. Si l'on considère l'espérance des membres de cette équation, il vient :

$$E \left[E(\varepsilon_t^4 | \varepsilon_{t-1}) \right] = E(\varepsilon_t^4)$$

$$3 E \left[E(\varepsilon_t^2 | \varepsilon_{t-1})^2 \right] \geq 3 \left[E E(\varepsilon_t^2 | \varepsilon_{t-1}) \right]^2 = 3 \left[E(\varepsilon_t^2) \right]^2$$

Ainsi, on en déduit que la loi marginale de ε_t a des queues plus épaisses qu'une loi normale puisque :

$$E(\varepsilon_t^4) \geq 3 \left[E(\varepsilon_t^2) \right]^2 \quad (3.44)$$

De plus, on peut calculer la kurtosis comme suit :

$$\begin{aligned}
Kurtosis &= \frac{E(\varepsilon_t^4)}{E(\varepsilon_t^2)^2} = \frac{3 \left[E(\varepsilon_t^2 | \varepsilon_{t-1}) \right]^2}{E(\varepsilon_t^2)^2} \\
&= 3 \frac{E(\varepsilon_t^2)^2}{E(\varepsilon_t^2)^2} + \frac{3}{E(\varepsilon_t^2)^2} \left\{ \left[E(\varepsilon_t^2 | \varepsilon_{t-1}) \right]^2 - E(\varepsilon_t^2)^2 \right\} \\
&= 3 + \frac{3}{E(\varepsilon_t^2)^2} \left\{ \left[E(\varepsilon_t^2 | \varepsilon_{t-1}) \right]^2 - E \left[E(\varepsilon_t^2 | \varepsilon_{t-1})^2 \right] \right\} \\
&= 3 + 3 \frac{Var \left[E(\varepsilon_t^2 | \varepsilon_{t-1}) \right]}{E(\varepsilon_t^2)^2}
\end{aligned}$$

La kurtosis est donc liée à une mesure de l'hétéroscédasticité conditionnelle.

Proposition 3.6. *Si le processus ε_t satisfait une représentation GARCH(p, q) conditionnellement gaussienne, telle que :*

$$\varepsilon_t = z_t \sqrt{h_t} \quad (3.45)$$

$$V(\varepsilon_t / \varepsilon_{t-1}) = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} = h_t \quad (3.46)$$

où z_t désigne bruit blanc faible gaussien, alors (i) la loi marginale de ε_t a des queues plus épaisses qu'une loi normale (distribution leptokurtique) :

$$E(\varepsilon_t^4) \geq 3 [E(\varepsilon_t^2)]^2 \quad (3.47)$$

et (ii) son coefficient d'excès de kurtosis peut s'exprimer sous la forme suivante :

$$\text{Excès de Kurtosis} = \frac{E(\varepsilon_t^4)}{E(\varepsilon_t^2)^2} - 3 = 3 \frac{Var \left[E(\varepsilon_t^2 | \varepsilon_{t-1}) \right]}{E(\varepsilon_t^2)^2} \quad (3.48)$$

Exemple 2 : On considère que un processus GARCH(1,1) tel que

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ N.i.d}(0, 1) \quad (3.49)$$

$$h_t = \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (3.50)$$

Bollerslev (1986) a montré d'une part que l'existence du moment d'ordre 4 de u exigeait $(\alpha_1 + \beta_1)^2 + 2\alpha_1^2 < 1$ et, d'autre part :

$$K_u = \frac{E(\varepsilon_t^4)}{E(\varepsilon_t^2)^2} = \frac{3 \left[1 - (\alpha_1 + \beta_1)^2 \right]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} \quad (3.51)$$

expression qui est toujours supérieure à 3 et peut donc aussi être utile sur des données à queues de distribution épaisses.

4. Estimation et Prévisions

Les modèles avec erreurs hétéroscédastiques peuvent être estimés généralement de trois façons :

- Estimateurs de la classe du Maximum de Vraisemblance (MV)
- Estimateurs du Pseudo Maximum de Vraisemblance (PMV)
- Estimateurs en deux étapes.

Nous ne présenterons ici que les deux premières méthodes d'estimation. La méthode en deux étapes (voir Bourbonnais et Terraza, 2004, pour un résumé) est asymptotiquement équivalente au MV, mais engendre à distance finie une perte d'efficacité. Commençons par étudier le cas des estimateurs du MV obtenue en cas d'erreur sur la spécification de la loi conditionnelle des résidus.

4.1. Estimateurs du MV sous l'hypothèse de normalité et Estimateurs du PMV

On reprend ici la présentation de Gouriéroux (1992). Soit un modèle tel que :

$$E\left(Y_t | \underline{Y}_{t-1}, X_t\right) = m_t\left(\underline{Y}_{t-1}, X_t, \theta\right) = m_t(\theta) \quad (4.1)$$

$$V\left(Y_t | \underline{Y}_{t-1}, X_t\right) = h_t\left(\underline{Y}_{t-1}, X_t, \theta\right) = h_t(\theta) \quad (4.2)$$

On note θ l'ensemble des paramètres intervenant à la fois dans l'expression de la moyenne conditionnelle et de la variance conditionnelle. La plupart des modèles ARCH que nous verrons peuvent être représentés sous cette forme. Nous commencerons par présenter les méthodes du MV et du PM, avant de présenter la procédure AUTOREG sous SAS.

4.1.1. Maximum et Pseudo Maximum de Vraisemblance appliqués aux modèles ARCH / GARCH

Nous présenterons de façon parallèle la méthode d'estimation du MV sous l'hypothèse de normalité de la distribution conditionnelle des résidus et la méthode d'estimation du PMV. En effet, l'idée générale des estimateurs du PMV consiste à démontrer que si l'on commet une erreur sur la distribution conditionnelle des résidus en utilisant à tort une log-vraisemblance fondée sur une loi normale, l'estimateur du MV ainsi obtenu peut tout de même être convergent si la vraie loi des résidus appartient à la même classe de loi que la loi normale (Gouriéroux, Montfort, 1989). L'estimateur sera (i) asymptotiquement convergent et (ii) asymptotiquement normal. Par conséquent la fonction de vraisemblance définissant l'estimateur du MV sous l'hypothèse de normalité et la fonction de pseudo-vraisemblance de l'estimateur du PMV sont les mêmes.

Definition 4.1. La fonction de log-vraisemblance associée à un échantillon de T observations (y_1, y_2, \dots, y_T) de Y_t sous l'hypothèse de normalité de la loi conditionnelle de Y_t sachant $\underline{Y_{t-1}}$ et X_t s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log[h_t(\theta)] - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{h_t(\theta)} \quad (4.3)$$

Attention, dans ces notations la fonction $h_t(\theta)$ désigne la variance conditionnelle et donc il n'y a pas lieu de l'élever au carré.

Exemple : Appliquons cette formule au cas d'un modèle de régression linéaire avec erreur ARCH(q) .

$$\begin{aligned} Y_t &= X_t \beta + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t(\theta)} \quad \text{avec } z_t \text{ N.i.d}(0, 1) \\ E(\varepsilon_t | \underline{\varepsilon_{t-1}}) &= 0 \quad V(\varepsilon_t | \underline{\varepsilon_{t-1}}) = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \end{aligned}$$

Dans cas on a donc :

$$\begin{aligned} E(Y_t | \underline{Y_{t-1}}, X_t) &= m_t(\theta) = X_t \beta \\ V(Y_t | \underline{Y_{t-1}}, X_t) &= h_t(\theta) = \alpha_0 + \sum_{i=1}^q \alpha_i (Y_{t-i} - \beta X_{t-i})^2 \end{aligned}$$

où $\theta = (\beta, \alpha_0, \alpha_1, \dots, \alpha_q) \in \mathbb{R}^{q+2}$. La log-vraisemblance s'écrit par conséquent :

$$\begin{aligned} \log L(\theta) &= -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log \left[\alpha_0 + \sum_{i=1}^q \alpha_i (Y_{t-i} - \beta X_{t-i})^2 \right] \\ &\quad - \frac{1}{2} \sum_{t=1}^T (y_t - X_t \beta)^2 \times \left[\alpha_0 + \sum_{i=1}^q \alpha_i (Y_{t-i} - \beta X_{t-i})^2 \right]^{-1} \end{aligned}$$

On peut déduire assez facilement l'écriture des CPO qui définissent l'estimateur du MV ou du PMV selon les cas.

Definition 4.2. Les estimateurs du MV sous l'hypothèse de normalité ou du PMV, notés $\hat{\theta}$, des paramètres $\theta \in \mathbb{R}^K$, satisfont un système non linéaire à K équations :

$$\left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0 \quad (4.4)$$

avec

$$\begin{aligned} \left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} &= -\frac{1}{2} \sum_{t=1}^T \frac{1}{h_t(\hat{\theta})} \left. \frac{\partial h_t(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} + \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\hat{\theta})]^2}{h_t(\hat{\theta})^2} \left. \frac{\partial h_t(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} \\ &\quad + \sum_{t=1}^T \left[\frac{y_t - m_t(\hat{\theta})}{h_t(\hat{\theta})} \right] \left. \frac{\partial m_t(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} \end{aligned}$$

Remarque : On peut montrer que ce système peut se décomposer en deux sous systèmes lorsque les paramètres θ interviennent de façon séparée dans l'écriture de l'espérance et de la variance conditionnelle. Ainsi, si l'on a $\theta = (\alpha \beta)'$ où α n'apparaît que dans l'espérance conditionnelle et β dans la variance conditionnelle, on peut décomposer ce système en deux sous système puisque :

$$\begin{aligned} \left. \frac{\partial \log L(\alpha)}{\partial \alpha} \right|_{\theta=\hat{\theta}} &= \sum_{t=1}^T \left[\frac{y_t - m_t(\hat{\alpha})}{h_t(\hat{\beta})} \right] \left. \frac{\partial m_t(\alpha)}{\partial \alpha} \right|_{\alpha=\hat{\alpha}} \\ \left. \frac{\partial \log L(\beta)}{\partial \beta} \right|_{\theta=\hat{\theta}} &= -\frac{1}{2} \sum_{t=1}^T \frac{1}{h_t(\hat{\beta})} \left. \frac{\partial h_t(\beta)}{\partial \beta} \right|_{\theta=\hat{\theta}} + \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\hat{\alpha})]^2}{h_t(\hat{\beta})^2} \left. \frac{\partial h_t(\beta)}{\partial \beta} \right|_{\beta=\hat{\beta}} \end{aligned}$$

Dans, le cas général du PMV, on sait que l'estimateur $\hat{\theta}$ est asymptotiquement normal et que sa matrice de variance covariance est définie par la formule suivante.

Definition 4.3. *Sous certaines conditions de régularité, l'estimateur du PMV est asymptotiquement convergent et normal.*

$$\sqrt{T} (\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1}) \quad (4.5)$$

où la matrice de variance covariance asymptotique de l'estimateur du PMV est calculée à partir de :

$$J = E_0 \left[-\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right] \quad I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right] \quad (4.6)$$

où E_0 désigne l'espérance prise par rapport à la vraie loi.

Naturellement dans la pratique les matrices I et J sont directement estimées en remplaçant l'espérance E_0 par la moyenne empirique et le paramètre inconnu θ par son estimateur convergent $\hat{\theta}$. Ainsi, on utilise :

$$\hat{I} = \frac{1}{T} \sum_{t=1}^T \left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} \left. \frac{\partial \log L(\theta)}{\partial \theta'} \right|_{\theta=\hat{\theta}} \quad (4.7)$$

$$\hat{J} = -\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\hat{\theta}} \quad (4.8)$$

et la variance estimée de $\hat{\theta}$ vérifie alors

$$\text{Var} \left[\sqrt{T} (\hat{\theta} - \theta) \right] = \hat{J}^{-1} \hat{I} \hat{J}^{-1} \quad (4.9)$$

Remarque 1 : Dans le cas où la vraie loi sous jacente est normale (Maximum de Vraisemblance), la matrice de variance covariance asymptotique se réduit à :

$$\text{Var} \left[\sqrt{T} (\hat{\theta} - \theta) \right] = J^{-1} \quad (4.10)$$

puisque $J = I$.

Remarque 2 : Dans le cas du MV lorsque l'on peut séparer les paramètres de l'espérance conditionnelle et de la variance conditionnelle, on montre que :

$$\text{Var} \left[\sqrt{T} (\hat{\beta} - \beta) \right] = \left[\frac{1}{T} \sum_{t=1}^T \frac{1}{2h_t(\hat{\beta})^2} \frac{\partial h_t(\beta)}{\partial \beta} \Big|_{\beta=\hat{\beta}} \frac{\partial h_t(\beta)}{\partial \beta'} \Big|_{\beta=\hat{\beta}} \right]^{-1} \quad (4.11)$$

4.1.2. La procédure AUTOREG : estimation par MV et PMV

Pour estimer un modèle de type ARCH / GARCH sous l'hypothèse de distribution conditionnelle normale par la méthode du maximum de vraisemblance (MV), on utilise la procédure AUTOREG. Considérons l'exemple suivant.

Exemple 1 : On cherche à estimer un modèle GARCH(1,1) sur données centrées à partir des données du SP500 présentées précédemment. Soit $dlsp_t$ le logarithme du rendement de l'indice Standard and Poor's 500.

$$dlsp_t = c + \varepsilon_t \quad (4.12)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ N.i.d}(0,1) \quad (4.13)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (4.14)$$

On utilise au préalable la procédure data de l'exemple donné dans le fichier example1.sas pour récupérer les données stockées dans le fichier "donnees". A partir de là, on utilise le programme suivant (fichier **example2.sas**) :

On obtient alors les résultats suivants :

L'estimation du modèle GARCH figure dans le bas de la sortie de résultats (figure 4.2). SAS donne le nom ARCH0 à la constante α_0 de l'équation

Figure 4.1: Estimation GARCH(1, 1) sur le rendement SP500

```

proc autoreg data=donnees;
    model disp = / garch=(q=1,p=1);
    output out=r r=yresid;
run;
quit;

```

de la variance conditionnelle, à ne pas confondre avec la constante c de l'équation de l'espérance conditionnelle. On obtient ainsi un modèle :

$$dlspt = \underset{(3.37)}{0.000465} + \varepsilon_t \quad (4.15)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ N.i.d}(0, 1) \quad (4.16)$$

$$h_t = \underset{(5.12)}{4.84e^{-7}} + \underset{(14.7)}{0.0443} \varepsilon_{t-1}^2 + \underset{(295)}{0.9519} h_{t-1} \quad (4.17)$$

Tous les paramètres de ce modèle sont largement significatifs. Par contre, on constate que le test de normalité de Jarque-Bera des résidus du modèle de variance conditionnelle z_t , conduit à rejeter l'hypothèse nulle de normalité au pour un seuil supérieur à 0.001. Donc, ce test semble indiquer que la variable z_t n'est pas distribuée selon une loi $N(0, 1)$. Admettons que z_t ne soit pas distribuée selon une loi normale : deux cas de figure sont envisageables.

- On peut tout d'abord suppose que la "vraie" distribution de z_t vérifie les conditions de régularité du PMV (Gouriéroux, Montfort et Trognon, 1989). L'estimateur dont la réalisation est sur la figure 4.2 correspond dans ce cas celui non plus à l'estimateur du MV, mais à celui du PMV. On sait qu'alors, cet estimateur est asymptotiquement convergent et normal. Toutefois, la formules permettant d'obtenir la matrice de variance covariance des paramètres doit être adaptée dans ce cas conformément à la définition 4.3. Par défaut, la procédure AUTOREG utilise une matrice de variance covariance définie par l'inverse de la matrice hessienne (\hat{J}^{-1} dans nos notations), alors que si l'on se situe dans le cas du PMV, la matrice de variance covariance asymptotique est définie par

$$\text{Var} \left[\sqrt{T} \left(\hat{\theta} - \theta \right) \right] = \hat{J}^{-1} \hat{I} \hat{J}^{-1}$$

Par conséquent, on doit spécifier l'option COVEST= QML dans le programme.

- La deuxième solution consiste à utiliser une autre distribution que la loi normale pour construire la vraisemblance. Naturellement, les formules de vraisemblance de la définition (4.1) ne sont plus valides. C'est que nous allons voir dans la prochaine section. Mais déjà signalons, que la procédure AUTOREG permet d'envisager en lieu et place d'une distribution normale sur z_t , une distribution de Student..

Figure 4.2: Résultats Estimation GARCH(1, 1)

```

Le Système SAS          10:09 Thursday, October
The AUTOREG Procedure
Dependent Variable     dlsp

Ordinary Least Squares Estimates

SSE          0.40117      DFE          3754
MSE          0.0001069   Root MSE     0.01034
SBC          -23672.063  AIC          -23678.294
Regress R-Square 0.0000      Total R-Square 0.0000
Durbin-Watson 2.0103

Variable      DDL      Estimation      Standard      t Value      Approx
              DDL      Estimation      Error         t Value      Pr > |t|

Intercept     1         0.000318        0.000169      1.88         0.0598

Algorithm converged.

GARCH Estimates

SSE          0.40125116   Observations 3755
MSE          0.0001069   Uncond Var   0.00012899
Log Likelihood 12286.4373      Total R-Square .
SBC          -24539.951   AIC          -24564.875
Normality Test 3199.6495      Pr > ChiSq   <.0001

Variable      DDL      Estimation      Standard      t Value      Approx
              DDL      Estimation      Error         t Value      Pr > |t|

Intercept     1         0.000465        0.000138      3.37         0.0008
ARCH0         1         4.8418E-7       9.4476E-8     5.12         <.0001
ARCH1         1         0.0443          0.003015      14.70        <.0001
GARCH1        1         0.9519          0.003226      295.09       <.0001

```

La syntaxe générale pour estimer un modèle de type ARCH / GARCH est la suivante :

PROC AUTOREG options ;

MODEL dependent = regressors / options ; (*Spécification* du modèle)

RESTRICT equation , ... , equation ; (Restrictions sur les paramètres)

TEST equation , ... , equation / option ; (*Test* d'hypothèse sur les paramètres)

OUTPUT OUT = SAS data set options ;

Voici quelques options de la commande MODEL qui permet de spécifier le modèle :

- CENTER : permet de centrer les variables et de supprimer la constante (quand le modèle n'a pas de regresseur)
- NOINT supprime la constante de la regression
- NLAG = ordre de la forme autorégressive Exemple : NLAG =2 ou NLAG = (1 3 6) si l'on spécifier une séquence de retards

- GARCH = (option-list) permet d'estimer un modèle ARCH / GARCH

Exemple 1 : model y = x1 x2 / garch=(q=1,p=1);

Exemple 2 : model y = x1 x2 / garch=(q=(1 3));

Exemple 3 : model y = / garch=(q=2,p=1,noint);

- Dans le premier cas, on estime un modèle GARCH(1,1) sur les résidus du modèle de regression linéaire de Y sur X_1, X_2 et une constante Dans le deuxième exemple, on estime un ARCH ($p = 0$) avec une structure ARCH à 1 et 3 retards (le coefficient du 2^{ème} retard est supposé nul). Dans le troisième exemple, on estime un modèle GARCH(2,1) directement sur la variable Y Attention dans ce cas, il n'y a pas de constante dans la variance conditionnelle, mais une constante dans l'espérance conditionnelle.
 - DIST= value permet de spécifier la distribution du terme d'erreur. On a que deux options
 - * DIST= T distribution de Student
 - * DIST= NORMAL pour une distribution normale (valeur par défaut)
 - NOINT : permet de ne pas inclure de constante dans l'équation de variance conditionnelle
 - NONNEG : contraint les parametres à être non négatifs
 - STATIONARY contraint les paramètres à satisfaire aux conditions de la stationnarité du 2nd ordre asymptotique (cf. IGARCH).
- COVEST= HESSIAN | QML, permet de spécifier la forme de la matrice de variance covariance asymptotique des paramètres. Sous l'hypothèse de normalité, cette matrice est construite de façon standard par inversion de la matrice hessienne (estimateur du MV). En revanche dans le cas du PMV, l'option QML (quasi maximum likelihood) permet de construire cette matrice conformément aux formules de la définition 4.3.

Exemple 2 : On souhaite estimer un modèle autorégressif d'ordre 2 sur le logarithme du rendement $dlspt_t$ de l'indice Standard and Poor's 500, avec une structure de type GARCH (1,3) sur la variance conditionnelle des résidus en imposant la nullité du paramètre associé à l'effet GARCH(2). On suppose que la distribution n'est pas normale, mais appartient à la famille des lois exponentielles (Gamma, Poisson etc..) et vérifie les conditions de régularité du PMV.

$$dlspt_t = c + \phi_1 dlspt_{t-1} + \phi_2 dlspt_{t-2} + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ i.i.d}(0, 1) \quad (4.18)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \beta_1 h_{t-1} \quad (4.19)$$

Sous cette hypothèse sur la distribution de z_t on doit construire un estimateur du PMV. Pour cela, on continue d'utiliser la procédure AUTOREG avec l'option par défaut DIST= NORMAL. On va donc maximiser la fonction de vraisemblance construite à tort sous l'hypothèse de normalité, mais on va adapter les formules de la matrice de variance covariance des paramètres en conséquence avec l'option COVEST= QML. Le programme est alors le suivant (fichier **Example2.sas**) :

Figure 4.3: Programme Estimation par PMV Modèle AR(2)-GARCH(3,1)

```
proc autoreg data=donnees;
  model dlsp = / nlag=2 garch=(q=(1 3),p=1) covest= qml ;
  output out=r r=yresid;
run;
```

Les résultats sont les suivants sont reportés sur la figure (4.4)

Figure 4.4: Estimation par PMV. Modèle AR(2)-GARCH(3,1) sur SP500

GARCH Estimates					
SSE		0.40156741	Observations		3755
MSE		0.0001069	Uncond Var		0.00011318
Log Likelihood		12292.177	Total R-Square		.
SBC		-24526.738	AIC		-24570.354
Normality Test		2395.3023	Pr > ChiSq		<.0001

Variable	DDL	Estimation	Standard Error	t Value	Approx Pr > t
Intercept	1	0.000506	0.000136	3.71	0.0002
AR1	1	-0.0166	0.0170	-0.98	0.3268
AR2	1	-0.005916	0.0199	-0.30	0.7665
ARCH0	1	2.8496E-7	3.4826E-7	0.82	0.4132
ARCH1	1	0.0729	0.0198	3.68	0.0002
ARCH3	1	-0.0436	0.0313	-1.39	0.1633
GARCH1	1	0.9682	0.0250	38.66	<.0001

4.1.3. La procédure AUTOREG : variances conditionnelles estimées et résidus

A partir de l'estimation d'un modèle ARCH/GARCH, la procédure AUTOREG permet de reconstruire assez facilement les variances conditionnelles estimées, les résidus et les résidus normalisés. La procédure AUTOREG autorise en effet un certain nombre d'outputs :

- **OUT**= nom du fichier de données dans lequel seront stockés les résultats sélectionnés.

- **CUSUM**= variable contenant les statistiques du Cusum. L'option CUSUMUB= (respectivement CUSUMULB) permet de spécifier le nom de la variable contenant les bornes supérieures (respectivement inférieures) de l'intervalle de confiance sur le cusum.
- **CUSUMSQ**= variable contenant les statistiques du Cusum squared (CUSUMSQUB et CUSUMSQLB pour les bornes).
- **CEV**= variable, (**HT**= variable), stocke dans la variable spécifiée la valeur de la variance conditionnelle du terme d'erreur $V(\varepsilon_t | \varepsilon_{t-1})$.
- **CPEV**= variable : writes the conditional prediction error variance to the output data set. The value of conditional prediction error variance is equal to that of the conditional error variance when there are no autoregressive parameters.
- **LCL**= name writes the lower confidence limit for the predicted value (specified in the PREDICTED= option) to the output data set. The size of the confidence interval is set by the ALPHACLI= option. When a GARCH model is estimated, the lower confidence limit is calculated assuming that the disturbances have homoscedastic conditional variance.
- **UCL**= name, writes the upper confidence limit for the predicted value (specified in the PREDICTED= option) to the output data set. The size of the confidence interval is set by the ALPHACLI= option. When the GARCH model is estimated, the upper confidence limit is calculated assuming that the disturbances have homoscedastic conditional variance.
- **LCLM**= name writes the lower confidence limit for the structural predicted value (specified in the PREDICTEDM= option) to the output data set under the name given. The size of the confidence interval is set by the ALPHACLM= option.
- **UCLM**= name writes the upper confidence limit for the structural predicted value (specified in the PREDICTEDM= option) to the output data set. The size of the confidence interval is set by the ALPHACLM=
- **PREDICTED**= name , P= name writes the predicted values to the output data set. These values are formed from both the structural and autoregressive parts of the model.
- **RESIDUAL**= name, R= name writes the residuals from the predicted values based on both the structural and time series parts of the model to the output data set.

Exemple 3 : Considérons l'exemple d'un modèle GARCH(1,1) sur le logarithme du rendement $dlspt$ de l'indice Standard and Poor's 500 sous l'hypothèse de normalité :

$$dlspt_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ N.i.d}(0, 1) \quad (4.20)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (4.21)$$

Le programme d'estimation avec les différentes options d'output est reproduit sur la figure (4.5) (fichier **Example3.sas**) :

Figure 4.5: Programme Estimation Modèle GARCH(1,1). Options de Sortie

```
proc autoreg data=donnees;
  model dlsp = / garch=(q=1,p=1);
  output out=resultsp r=epsilon cev=condvar p=prevision alphacli=0.05
    lcl=lower ucl=upper;
run;
```

Dans ce cas, on a spécifié différentes options de sortie. Les variables de sortie sont stockées dans le fichier resultsp.

- Dans la variable epsilon figurent les résidus estimés $\hat{\varepsilon}_t$.
- La variable condvar correspond aux valeurs estimées de la variance conditionnelle \hat{h}_t .
- La prévision associée à cette variable est stockée dans la variable prevision, ainsi que les bornes inférieures et supérieures au seuil de 5%

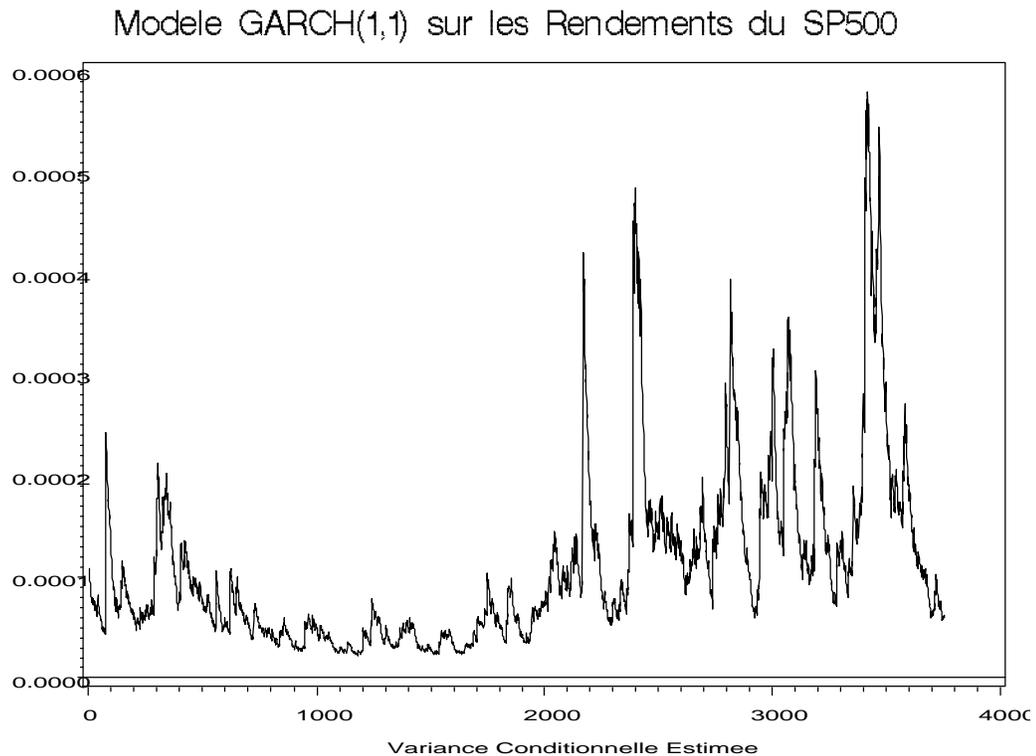
Supposons que l'on veuille faire le graphique de la variance conditionnelle estimée. Le programme correspondant peut s'écrire sous la forme suivante (fichier **Example3.sas**) :

```
proc gplot data = resultsp;
  plot condvar*t / vref=0 haxis=axis1 vaxis=axis2;
  symbol1 i=join v=none;
  axis1 label=( h=1 'Variance Conditionnelle Estimee');
  axis2 label=none;
run;
```

Figure 4.6:

Le graphique de la variance conditionnelle estimée est alors obtenu sur la figure (4.7).

Figure 4.7: Variance Conditionnelle Estimée. Modèle GARCH(1,1)



Il peut être en outre intéressant de grapher sur une même figure, les estimations de la variance conditionnelle et des rendements du SP500. Le programme peut être alors le suivant (fichier **Example3.sas**) :

Le résultat de ce programme est reporté sur la figure (4.9). On voit bien dès lors coïncider l'augmentation de la variance conditionnelle sur la période 2000-2003 et ce en même temps que l'on observe sur les rendements de forts clusters de volatilité très rapprochés.

Figure 4.8: Programme Graphique de la Variance Conditionnelle Estimée et des Rendements

```

proc gplot data = resultsp;
  plot dlsp*t=1 / vref=0 haxis=axis4 vaxis=axis2;
  plot2 condvar*t=2 / vref=0 haxis=axis2 vaxis=axis2;
  symbol1 i=join v=none;
  symbol2 i=join v=none c=red;
  axis4 label=(h=1.5 c=black 'Rendements SP500 et ' c=red 'Variance Conditionnell
run;

```

4.1.4. La procédure MODEL

Il existe une autre façon d'estimer des modèles ARCH / GARCH qui autorise plus de libertés quant au choix de la distribution des résidus : il s'agit de la procédure MODEL. Cette procédure permet d'estimer un modèle non linéaire de la forme :

$$y = f(x, \theta) + \varepsilon$$

Cette procédure permet entre autres d'estimer le modèle par les MCO et Double MCO, les méthodes SUR et SUR itératif, les Triple Moindres Carrés, les GMM et le maximum de vraisemblance à information complète ou FIML. C'est précisément cette dernière possibilité qui nous intéresse dans le cas des modèles ARCH/ GARCH. La syntaxe générale est de la forme suivante (entres autres) :

```

PROC MODEL options;
ENDOGENOUS variable [ initial values ] ... ;
ESTIMATE item [ , item ... ] [ ,/ options ] ;
EXOGENOUS variable [ initial values ] ... ;
OUTVARS variable ... ;
PARAMETERS variable [ value ] variable [ value ] ... ;
SOLVE variables [SATISFY=(equations) ] [ / options ] ;
TEST [ "name" ] test1 [ , test2 ... ] [ / options ] ;
VAR variable [ initial values ] ... ;

```

Exemple : On cherche à estimer un modèle GARCH(1,1) sous l'hypothèse de normalité sur le rendement du SP500.

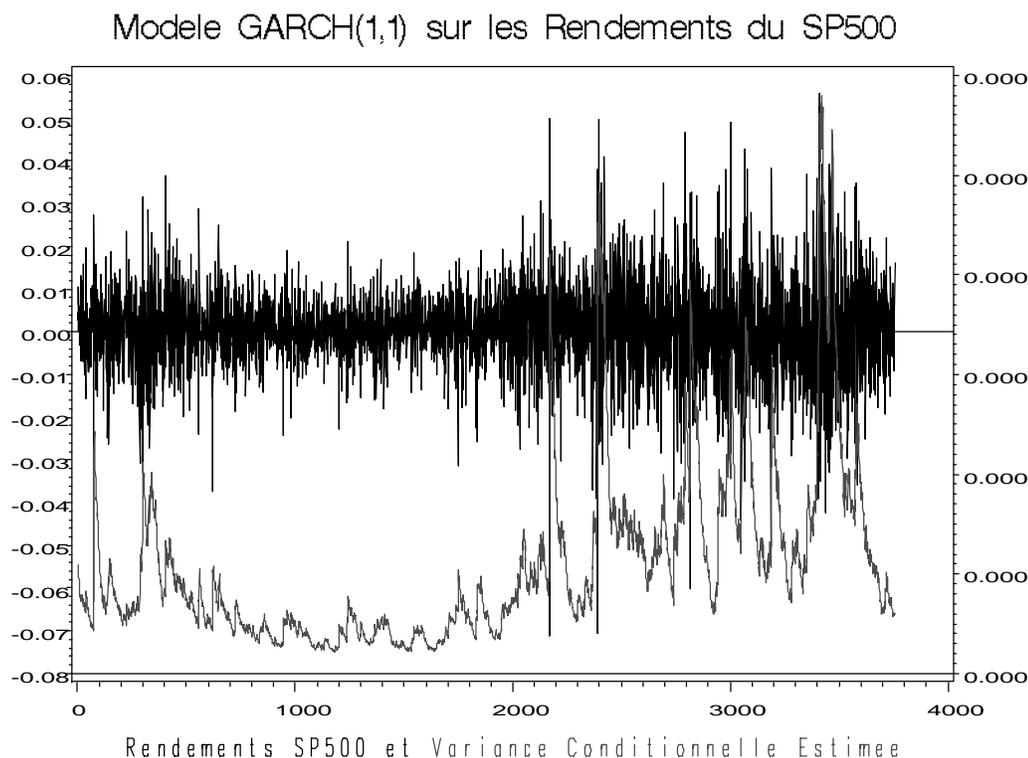
$$dlsp_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ N.i.d. } (0, 1) \tag{4.22}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{4.23}$$

Le programme d'estimation de ce modèle par la procédure MODEL est reporté sur la figure 4.10. Dans cette procédure, il est nécessaire d'indiquer dans l'option FIT que l'on utilise la méthode du maximum de vraisemblance à information complète (FIML).

Figure 4.9: Variance Conditionnelle Estimée et Rendements. Modèle GARCH(1,1)



Le résultat de cette estimation est reporté sur la figure 4.11.

Le résultat de l'estimation de ce modèle GARCH(1,1) par la procédure AUTOREG sous l'hypothèse de normalité est reporté sur le graphique 4.12. On observe que les deux résultats d'estimation sont proches mais pas totalement similaires. On peut obtenir un tel résultat en contrôlant les différentes options de ces deux procédures (cf. exercice 3).

Dans la procédure MODEL, on peut entre autres contrôler les options relatives au choix des conditions initiales de l'algorithme d'optimisation, par l'option PARAM en listant les paramètres du modèle et en leur associant une valeur initiale. On peut en outre réaliser l'optimisation de la fonction de vraisemblance sous certaines contraintes (non négativité par exemple) sur les paramètres via l'option BOUNDS. Dans le programme () on contraint ainsi les estimateurs des paramètres α_0 , α_1 et β_1 à être positifs ou nuls. Mais l'essentiel des options figure dans la commande FIT qui permet de déterminer toutes les options de la méthode d'estimation et de l'algorithme d'optimisation de la vraisemblance dans notre cas. Les principales options sont les suivantes :

- **PARMS=** permet de sélectionner un sous ensemble de paramètre sur lesquels porte l'optimisation.

Figure 4.10: Estimation d'un Modèle GARCH(1,1) par la procédure MODEL

```

proc model data=donnees;
  /* Equation Esperance Conditionnelle */
  dlsp= intercept ;
  /* Variance Conditionnelle */
  h.dlsp = arch0 + arch1 * zlag(resid.dlsp **2 ) +
           garch1 * zlag(h.dlsp) ;
  fit dlsp / fiml ;
run;

```

- **DROP=** (parameters ...) permet de spécifier le nom des paramètres à ne pas estimer.
- **COVBEST=GLS | CROSS | FDA** specifies the variance-covariance estimator used for FIML. COVBEST=GLS selects the generalized least-squares estimator. COVBEST=CROSS selects the crossproducts estimator. COVBEST=FDA selects the inverse of the finite difference approximation to the Hessian. The default is COVBEST=CROSS.
- **FIML** specifies full information maximum likelihood estimation.
- **VARDEF=N | WGT | DF | WDF** specifies the denominator to be used in computing variances and covariances. VARDEF=N specifies that the number of nonmissing observations be used. VARDEF=WGT specifies that the sum of the weights be used. VARDEF=DF specifies that the number of nonmissing observations minus the model degrees of freedom (number of parameters) be used. VARDEF=WDF specifies that the sum of the weights minus the model degrees of freedom be used. The default is VARDEF=DF. VARDEF=N is used for FIML estimation.
- **OUT= SAS-data-set** names the SAS data set to contain the residuals, predicted values, or actual values from each estimation. Only the residuals are output by default.
- **OUTCOV ou COVOUT** writes the covariance matrix of the estimates to the OUTEST= data set in addition to the parameter estimates. The OUTCOV option is applicable only if the OUTEST= option is also specified
- **OUTPREDICT** writes the predicted values to the OUT= data set. This option is applicable only if OUT= is specified.
- **OUTRESID** writes the residual values computed from the parameter estimates to the OUT= data set. The OUTRESID option is the default if neither OUT-

Figure 4.11: Estimation d'un Modèle GARCH(1,1) par la procédure MODEL

The MODEL Procedure							
Nonlinear FIML Summary of Residual Errors							
Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq
d1sp	4	3751	0.4013	0.000107	0.0103	-0.0004	-0.0012
resid.d1sp		3751	3751.7	1.0002	1.0001		
Nonlinear FIML Parameter Estimates							
Parameter	Estimation	Approx Std Err	t Value	Approx Pr > t			
intercept	0.000533	0.000132	4.05	<.0001			
arch0	1.204E-6	2.25E-7	5.35	<.0001			
arch1	0.065006	0.00782	8.31	<.0001			
garch1	0.925165	0.00843	109.79	<.0001			
Number of Observations			Statistics for System				
Used	3755	Log Likelihood	12278				
Missing	1						

PREDICT nor OUTACTUAL is specified. This option is applicable only if the OUT= option is specified.

- **NORMAL** performs tests of normality of the model residuals.

Les principales options qui permettent de contrôler l'algorithme d'optimisation sont les suivantes :

- **ITALL** specifies all iteration printing-control options (I, ITDETAILS, ITPRINT, and XPX). ITALL also prints the crossproducts matrix (labeled CROSS), the parameter change vector, and the estimate of the cross-equation covariance of residuals matrix at each iteration.
- **ITPRINT** prints the parameter estimates, objective function value, and convergence criteria at each iteration.
- **CONVERGE= value1** ou **CONVERGE= (value1, value2)** specifies the convergence criteria. The convergence measure must be less than value1 before convergence is assumed. value2 is the convergence criterion for the S and V matrices for S and V iterated methods. value2 defaults to value1. See "The Convergence Criteria" for details. The default value is CONVERGE=.001.
- **HESSIAN= CROSS | GLS | FDA** specifies the Hessian approximation used for FIML. HESSIAN=CROSS selects the crossproducts approximation to the

Figure 4.12: Estimation d'un Modèle GARCH(1,1) par la procédure AUTOREG

```

The AUTOREG Procedure
Dependent Variable   disp

Ordinary Least Squares Estimates

SSE           0.40117      DFE           3754
MSE           0.0001069   Root MSE     0.01034
SBC           -23672.063   AIC          -23678.294
Regress R-Square 0.0000      Total R-Square 0.0000
Durbin-Watson  2.0103

Variable      DDL      Estimation      Standard      t Value      Approx
              DDL      Estimation      Error         Pr > |t|
Intercept     1        0.000318        0.000169     1.88         0.0598

Algorithm converged.

GARCH Estimates

SSE           0.40125116   Observations  3755
MSE           0.0001069   Uncond Var   0.00012899
Log Likelihood 12286.4373      Total R-Square .
SBC           -24539.951   AIC          -24564.875
Normality Test 3199.6495      Pr > ChiSq   <.0001

Variable      DDL      Estimation      Standard      t Value      Approx
              DDL      Estimation      Error         Pr > |t|
Intercept     1        0.000465        0.000138     3.37         0.0008
ARCH0         1        4.8418E-7       9.4476E-8    5.12         <.0001
ARCH1         1        0.0443          0.003015     14.70        <.0001
GARCH1        1        0.9519          0.003226     295.09       <.0001

```

Hessian, HESSIAN=GLS selects the generalized least-squares approximation to the Hessian, and HESSIAN=FDA selects the finite difference approximation to the Hessian. HESSIAN=GLS is the default.

- **MAXITER= n** specifies the maximum number of iterations allowed. The default is MAXITER=100.
- **METHOD= GAUSS | MARQUARDT** specifies the iterative minimization method to use. METHOD=GAUSS specifies the Gauss-Newton method, and METHOD=MARQUARDT specifies the Marquardt-Levenberg method. The default is METHOD=GAUSS. See "Minimization Methods" for details.

Exemple : on reprend l'estimation du modèle GARCH(1,1) de l'exemple précédent en utilisant différentes options de la procédure MODEL (**example4.sas**).

Figure 4.13: Estimation d'un Modèle GARCH(1,1) par la procédure MODEL avec Options

```

proc model data=donnees;
  parameters arch0 .0005 arch1 .0443 garch1 .96 ;
  /* Equation Esperance Conditionnelle */
  dlsp= intercept ;
  /* Variance Conditionnelle */
  h.dlsp = arch0 + arch1 * zlag(resid.dlsp * resid.dlsp ) +
           garch1 * zlag(h.dlsp) ;
  bounds arch0 arch1 garch1 >= 0;
  fit dlsp/ method=marquardt fiml itprint normal
        hessian=cross CONVERGE=.00001 out=resultats;
  outvars resid.dlsp h.dlsp t;
run;

```

4.2. Estimateurs du MV sous d'autres lois

Comme nous l'avons mentionné, on peut envisager d'autres lois que la loi normale dans la procédure d'estimation par MV. On considère un modèle où les résidus sont définis par la quantité

$$\varepsilon_t = z_t \sqrt{h_t} \quad (4.24)$$

. Trois lois de distribution sont parfois imposées sur l'aléatoire z_t en dehors de la loi normale : Student, skewed-student et GED. La seconde et la quatrième sont des lois symétriques qui ont des queues de distribution plus épaisses que la normale et donc à même de prendre en compte ce phénomène souvent observé sur les séries économiques, notamment lorsqu'il s'agit de grandeurs financières à fréquence d'observation élevée. La troisième est dissymétrique et peut donc prétendre à la modélisation de séries ayant une skewness non nulle. On rappelle brièvement chacune de ces lois et pour finir la valeur du terme $E[|z_t|]$ qui apparaît dans le modèle EGARCH.

Nous commencerons par présenter ces différentes lois.

4.2.1. La distribution de Student

La distribution de Student est souvent utilisée pour prendre en compte un éventuel effet de kurtosis.

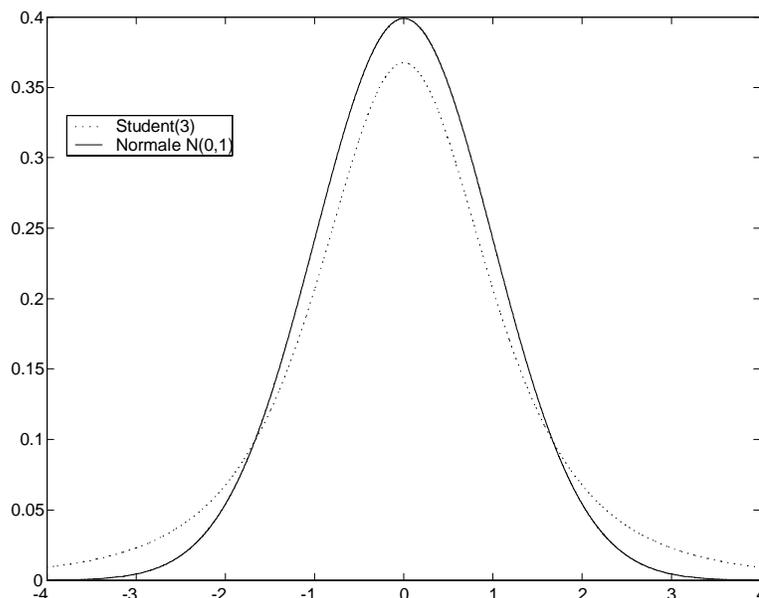
Rappel Si x et y sont deux variables aléatoires indépendantes, telles que x suit une loi $N(0, 1)$ et y suit une loi du chi-deux à v degrés de liberté, alors la variable

$$t = \frac{x}{\sqrt{y/v}}$$

est distribuée selon une loi de Student à v degrés de liberté, telle que $E(t) = 0$ si $v > 1$ et $V(t) = v/(v-2)$ si $v > 2$.

Sur le graphique 4.14, sont reportées les densités d'une loi normal et d'une loi de Student à 3 degrés de libertés. On vérifie que cette dernière admet des queues de distribution plus épaisses que celles de la loi normal : pour des degrés de liberté faibles, la distribution de Student est donc une distribution leptokurtique.

Figure 4.14: Comparaison entre les Distributions de Student et Normale



On admettra le résultat suivant.

Definition 4.4. Si la variable z_t admet une distribution de Student à v degrés de libertés, où $v \in \mathbb{N}$ vérifie $v > 2$, alors la log-vraisemblance associée à une observation z_t et à l'ensemble de paramètres θ s'écrit :

$$\log L(\theta, \varepsilon_t) = \log \left[\Gamma \left(\frac{v+1}{2} \right) \right] - \log \left[\Gamma \left(\frac{v}{2} \right) \right] - 0.5 \left\{ \left[\log [\pi (v-2)] + \log (h_t) + (1+v) \log \left(1 + \frac{z_t^2}{v-2} \right) \right] \right\} \quad (4.25)$$

où $\Gamma(\cdot)$ désigne la fonction Gamma.

On rappelle que :

$$\Gamma(r) = \int_0^\infty e^{-x} x^{r-1} dx \quad r > 0$$

et que

$$\Gamma(\alpha) = (\alpha - 1)! \quad \text{si } \alpha \in \mathbb{N}^*$$

4.2.2. La distribution de Student dissymétrique standardisée

Elle est introduite dans le cadre des divers processus GARCH par Lambert et Laurent (2001) qui se fondent sur une procédure de Fernández et Steel (1998) permettant d'introduire de la skewness dans toute distribution unimodale et symétrique autour de zéro. Ils l'appliquent à la loi de Student pour définir la Student dissymétrique qu'ils standardisent afin d'obtenir une densité écrite en fonction de l'espérance et de variance de l'aléatoire³. La log-vraisemblance correspondance est alors :

$$l(\theta, y_t) = \log\left(\Gamma\left(\frac{v+1}{2}\right)\right) - \log\left(\Gamma\left(\frac{v}{2}\right)\right) + \log\left(\frac{2}{\xi + \frac{1}{\xi}}\right) + \log(s) \quad (4.26)$$

$$- 0.5 \left[\log(\pi(v-2)) + \log(h_t) + (1+v) \log\left(1 + \frac{(sz_t + m)^2}{v-2} \xi^{-2I_t}\right) \right]$$

avec :

$$m = \frac{\Gamma\left(\frac{v-1}{2}\right) \sqrt{v-2}}{\sqrt{\pi} \Gamma\left(\frac{v}{2}\right)} \left(\xi - \frac{1}{\xi}\right) \quad (4.27)$$

$$s^2 = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2 \quad (4.28)$$

$$I_t = \begin{cases} 1 & \text{si } z_t \geq -\frac{m}{s} \\ 0 & \text{si } z_t < -\frac{m}{s} \end{cases} \quad (4.29)$$

ξ est un indicateur de dissymétrie tel que lorsque $\xi = 1$, la distribution de Student dissymétrique standardisée est égale à la distribution de Student précédente.

4.2.3. La distribution Generalized Error Distribution

La distribution Generalized Error Distribution (GED) est définie par :

Definition 4.5. Si la variable z_t , telle que $E(z_t) = 0$ et $Var(z_t) = 1$, admet une distribution GED de paramètre $v > 0$, sa densité est définie par :

$$f_z(z_t) = \frac{v \exp[-(1/2)|z_t/\lambda|^v]}{\lambda 2^{[(v+1)/v]} \Gamma(1/v)} \quad (4.30)$$

où $\Gamma(\cdot)$ désigne la fonction gamma et λ est une constante définie par :

$$\lambda = \left\{ \frac{2^{-\frac{2}{v}} \Gamma(1/v)}{\Gamma(3/v)} \right\}^{\frac{1}{2}} \quad (4.31)$$

³La student dissymétrique est en effet définie sur un mode (qui n'est pas l'espérance) et une mesure de volatilité (qui n'est pas la variance) conditionnels. Pour plus de détails, voir Lambert et Laurent (2001).

Le paramètre v détermine l'épaisseur des queues de distribution. Si $v = 2$, alors $\lambda = 1$ et l'on retrouve la densité d'une loi normale $N(0, 1)$. Si $v < 2$, les queues de distribution sont plus épaisses que celles d'une loi normale (distribution leptokurtique). Si $v > 2$, la distribution est platykurtique. Pour cette raison, elle est souvent utilisée afin de prendre en compte des effets de kurtosis. On note en particulier que :

$$E|z_t| = \frac{\lambda 2^{1/v} \Gamma(2/v)}{\Gamma(1/v)} \quad (4.32)$$

Préconisée notamment par Nelson (1991), la log-vraisemblance associée à une distribution de type Generalized Error Distribution (GED) est la suivante :

Definition 4.6. *Si la variable z_t admet une distribution GED avec $v \in \mathbb{R}^*$, alors la log-vraisemblance associée à une observation z_t et à l'ensemble de paramètres θ s'écrit :*

$$l(\theta, z_t) = \log(v/\lambda) - 0.5 \left| \frac{z_t}{\lambda} \right|^v - (1 + v^{-1}) \log(2) - \log \left[\Gamma \left(\frac{1}{v} \right) \right] - 0.5 \log(h_t) \quad (4.33)$$

avec :

$$\lambda = \sqrt{\frac{2^{-\frac{2}{v}} \Gamma(\frac{1}{v})}{\Gamma(\frac{3}{v})}} \quad (4.34)$$

où $\Gamma(\cdot)$ désigne la fonction Gamma.

4.2.4. La procédure AUTOREG

Sous SAS, la procédure AUTOREG n'autorise en option que deux distributions dans la procédure de MV : la distribution normale (par défaut) et la distribution de Student.. Par exemple, si l'on souhaite estimer le modèle suivant

$$\begin{aligned} dlsp_t &= c + \phi_1 dlsp_{t-1} + \phi_2 dlsp_{t-2} + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t} \end{aligned} \quad (4.35)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \beta_1 h_{t-1} \quad (4.36)$$

en supposant que la variable z_t suit une distribution de Student à v degrés de liberté, on peut utiliser le programme suivant (fichier **Example5.sas**) :

Figure 4.15: Estimation d'un AR(2)-GARCH(3,1) sous l'hypothèse d'une distribution de Student

```
proc autoreg data=donnees;
  model dlsp = / nlag=2 garch=(q=(1 3),p=1) dist= t;
  output out=r r=yresid;
run;
```

Figure 4.16: Estimation d'un AR(2)-GARCH(3,1) sous l'hypothèse d'une distribution de Student

GARCH Estimates					
SSE		0.40122616	Observations		3755
MSE		0.0001069	Uncond Var		.
Log Likelihood		14457.071	Total R-Square		.
SBC		-28848.295	AIC		-28898.142
Normality Test		5752.8219	Pr > ChiSq		<.0001

Variable	DDL	Estimation	Standard Error	t Value	Approx Pr > t
Intercept	1	0.000520	0.000101	5.12	<.0001
AR1	1	0.005989	0.0142	0.42	0.6740
AR2	1	0.0182	0.0140	1.30	0.1921
ARCH0	1	9.2511E-6	0.0000282	0.33	0.7426
ARCH1	1	0.7571	2.2852	0.33	0.7404
ARCH3	1	1.1003	3.3777	0.33	0.7446
GARCH1	1	0.8911	0.0167	53.47	<.0001
TDF1	1	0.4869	0.0399	12.21	<.0001

Les résultats sont reportés sur la figure (4.16)

Les résultats de l'estimation par MV sous l'hypothèse d'une distribution de Student sont sensiblement différents de ceux obtenus sous l'hypothèse de normalité (figure 4.4), ce qui met en évidence l'influence du choix de la distribution dans la construction des estimateurs des paramètres de la variance conditionnelle et de l'espérance conditionnelle. En plus, des paramètres obtenus dans le cas normal (figure 4.4), SAS fournit un estimateur non pas du nombre de degré de liberté du Student, mais de son inverse.

$$\hat{v} = \frac{1}{0.4869} = 2.0538 \quad (4.37)$$

On constate ici que le nombre de degré de libertés estimé est compatible avec l'existence du moment d'ordre deux de la variable z_t .

4.2.5. La procédure MODEL

La procédure MODEL, plus souple que la procédure AUTOREG, permet d'estimer les modèles ARCH-GARCH sous de nombreuses distributions et plus uniquement sous la distribution de Student. Nous estimerons un modèle GARCH sous l'hypothèse de normalité, puis sous l'hypothèse de distribution GED.

Exemple 1 (fichier exemple5.sas): On cherche à estimer le modèle suivant

$$dls p_t = c + \phi_1 dls p_{t-1} + \varepsilon_t \quad (4.38)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (4.39)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (4.40)$$

en supposant que la variable z_t suit une distribution de Student à v degrés de liberté. Pour cela on utilise la procédure MODEL comme suit (figure 4.17) :

Figure 4.17: Estimation d'un AR(1)-GARCH(1,1) sous l'hypothèse d'une distribution de Student. Procédure MODEL.

```
proc model data=donnees;
  parameters ar1 .005 arch0 1.00 arch1 0.5 garch1 .5 df 5;
  dlsp = intercept + ar1*lag(dlsp);
  h.dlsp = arch0+arch1*xlag(resid.dlsp **2,mse.dlsp)
          + garch1*xlag(h.dlsp,mse.dlsp);
  errormodel dlsp ~ t(h.dlsp,df);
  fit dlsp / fiml method=marquardt out=resultat itprint ;
  outvars resid_dlsp h_dlsp t;
run;
```

Le programme est identique à ceux présentés dans le cas normal, à l'exception de la ligne ERRORMODEL qui permet de spécifier un certain nombre de distribution "pré-programmées" pour les résidus z_t . Les résultats d'estimation sont reportés sur la figure 4.18. Contrairement au cas de la procédure AUTOREG, cette fois-ci le degré de liberté de la loi de Student est estimé directement (paramètre DF).

Figure 4.18: Estimation d'un AR(1)-GARCH(1,1) sous l'hypothèse d'une distribution de Student. Procédure MODEL.

The MODEL Procedure							
Nonlinear Likelihood Summary of Residual Errors							
Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq
dlsp	6	3748	0.4014	0.000107	0.0103	-0.0005	-0.0018
NRESID.dlsp		3748	5766.7	1.5386	1.2404		
Nonlinear Likelihood Parameter Estimates							
Parameter	Estimation	Approx Std Err	t Value	Approx Pr > t			
ar1	-0.00132	0.0163	-0.08	0.9359			
arch0	1.914E-7	9.545E-8	2.00	0.0450			
arch1	0.031337	0.00518	6.05	<.0001			
garch1	0.951274	0.00761	125.05	<.0001			
df	5.777288	0.5860	9.86	<.0001			
intercept	0.000544	0.000119	4.56	<.0001			
Number of Observations				Statistics for System			
Used	3754		Log Likelihood	12416			
Missing	1						

Le problème c'est qu'il existe un nombre relativement limité de distribution disponibles avec l'option ERRORMODEL. En effet, on ne peut considérer que les lois suivantes :

- Lois de Cauchy : CAUCHY(<location, scale>)
- Loi du Chi-Deux : CHISQUARED (df < , nc>)
- Loi de Poisson : POISSON(mean)
- Loi de Student : T(v1 v2 ... vn, df)
- Loi Uniforme : UNIFORM(<left, right>)
- Loi de Fisher F(ndf, ddf < , nc>)

Donc on peut être amené à écrire directement la densité de la loi de distribution des termes d'erreur z_t . Envisageons le cas où l'on cherche à estimer un modèle ARCH-GARCH sous une distribution de type GED. Pour cela on utilise l'option GENERAL(Likelihood, <parm1, parm2, ... parm n >)

Exemple 2 (fichier example6.sas): On cherche à estimer le modèle suivant

$$dlsp_t = c + \varepsilon_t \quad (4.41)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (4.42)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (4.43)$$

en supposant que la variable z_t suit une distribution GED (Nelson, 1991). Le programme d'estimation est alors le suivant (figure 4.19) :

Figure 4.19: Estimation d'un GARCH(1,1) sous Distribution GED. Procedure MODEL.

```
proc model data = donnees;
  parms nu 2 arch0 .1 arch1 .2 garch1 .75;
  /* Esperance Conditionnelle */
  dlsp = intercept ;
  /* Variance Conditionnelle */
  h.dlsp = arch0 + arch1 * xlag(resid.dlsp ** 2, mse.dlsp) +
    garch1 * xlag(h.dlsp, mse.dlsp);
  /* Distribution des Erreurs */
  lambda = sqrt(2**(-2/nu) * gamma(1/nu) / gamma(3/nu)) ;
  obj = log(nu/lambda) - (1 + 1/nu) * log(2) - lgamma(1/nu) -
    .5*abs(resid.dlsp/lambda/sqrt(h.dlsp))**nu - .5*log(h.dlsp) ;
  obj = -obj ;
  errormodel dlsp ~ general(obj,nu);
  /* estimation */
  fit dlsp / method=marquardt out=resultat itprint fml ;
  outvars resid_dlsp h_dlsp t;
run;
```

On reconnaît dans la variable obj la valeur de l'opposé de la log-vraisemblance associée à une observation z_t .

4.3. Prévisions et intervalles de confiance

Comme nous l'avons dit précédemment, une des utilisations possibles des estimations de la variance conditionnelle consiste à proposer des intervalles de confiance sur la variable endogène construits sans supposer l'invariance dans le temps des moments d'ordre deux. Ainsi, comme l'indique Gouriéroux (1992), la principale différence entre les modélisations ARMA et ARCH apparaît lorsque l'on compare leurs intervalles de confiance. Même lorsque les deux modèles conduisent à des valeurs ajustées proches, d'un côté avec la modélisation ARMA, l'intervalle de confiance est construit avec une variance constante dans le temps, ce qui n'est pas le cas avec un modèle présentant une structure ARCH/GARCH sur les résidus.

Estimons un modèle autorégressif sur le taux de croissance de l'indice des prix à la consommation (variable CPI, **Example7_CPI.sas**), sur données mensuelles (janvier 1980-décembre 2002) à partir d'un modèle AR(3) avec résidus GARCH(1,1) sous distribution de Student.. Le programme d'estimation et d'affichage des résultats sous forme de graphique est le suivant :

Figure 4.20: Prévision et Intervalle de Confiance sur CPI. Modèle AR(3)-GARCH(1,1)

```
proc autoreg data=atacpi;
  model dcpi = / nlag=3 garch=(p=1,q=1) dist= t;
  output out=resgarch r=epsilon p=prevision alphacli=0.05
         lcl=lowerarch ucl=upperarch;
run;



---


title ' Modele GARCH avec Distribution de Student';



---


proc gplot data=resgarch;
  plot dcpi*t=1 prevision*t=2 lowerarch*t=3
       upperarch*t=3 / overlay ;
  symbol1 v=none i=join;
  symbol2 v=circle i=join;
  symbol3 v=none i=join;
run;
```

Les résultats de ce programme sont reproduits sur la figure (

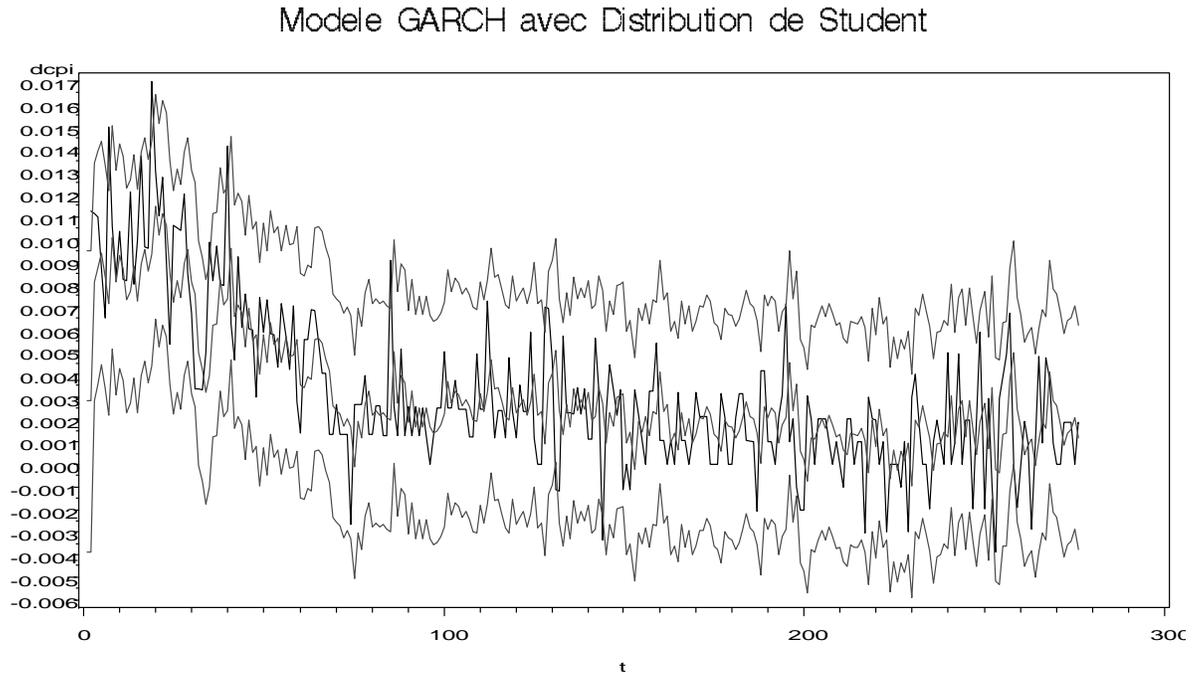
Il y a un Problème : reprendre avec Charpentier + estimation et prévision

4.4. Tests d'effets ARCH / GARCH

Comment tester la présence des effets ARCH dans la série Y_t ou dans le résidu du modèle linéaire auto-régressif ? Deux principaux tests existent :

1. Tests d'autocorrélation sur les carrés ε_t^2 : application des statistiques usuelles du type Q stat (Box Pierce, Ljung Box etc..)

Figure 4.21: Prédiction et Intervalle de Confiance sur CPI. Modèle AR(3)-GARCH(1,1)



2. Tests LM d'absence d'auto-corrélation sur ε_t^2

On rappelle que si l'on note r_k l'autocorrélation d'ordre k d'un processus $\{\varepsilon_t^2, t \in \mathbb{Z}\}$. Pour un ordre K , le test de Box et Pierce est le test de l'hypothèse $H_0 : r_1 = \dots = r_K = 0$ contre $H_1 : \exists j \in [1, K]$, tel que $r_j \neq 0$. C'est donc un test de nullité des K premières autocorrélations du processus considéré. Pour un processus $ARMA(p, q)$, la statistique de ce test est :

$$Q_{BP}(K) = T \sum_{k=1}^K r_k^2 \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \chi^2(K - p - q)$$

L'hypothèse H_0 est rejetée au seuil de 5% si Q_{BP} est supérieur au quantile 0.95 de la loi du χ^2 correspondant. Dans le cas du tests de Ljung-Box, ces statistiques, définies pour un ordre K , correspondent à l'hypothèse nulle $H_0 : r_k = 0 \forall k \leq K$ et sont construites de la façon suivante :

$$Q(K) = T(T+2) \sum_{k=1}^K \frac{r_k^2}{T-k} \xrightarrow[T \rightarrow \infty]{\mathcal{L}} \chi^2(K - p - q)$$

En ce qui concerne la statistique de test $LM(K)$, il s'agit tout simplement de la statistique associée au test de l'hypothèse $H_0 : \beta_1 = \beta_2 = \dots = \beta_K = 0$ dans le modèle

$$\widehat{\varepsilon}_t^2 = \beta_0 + \beta_1^2 \widehat{\varepsilon}_{t-1}^2 + \dots + \beta_{t-K}^2 \widehat{\varepsilon}_{t-K}^2 + \mu_t \quad (4.44)$$

où $\widehat{\varepsilon}_{t-j}^2$ désigne le résidu estimé au carré d'un modèle d'espérance conditionnelle. On rappelle en outre le principe du test du score ou du multiplicateur de Lagrange. On sait que si l'hypothèse nulle est satisfaite, les deux estimateurs non contraint $\widehat{\beta}_j$ et contraint $\widehat{\beta}_j^c$ doivent relativement proches l'un de l'autre, et que donc la même propriété doit être vérifiée pour le vecteur des conditions du premier ordre de la maximisation de la log vraisemblance..

Definition 4.7. *La statistique LM_j du test du multiplicateur de Lagrange associée au test unidirectionnel $H_0 : \beta = a$, où $a \in R^K$ contre $H_1 : \beta \neq a$ admet la loi suivante sous H_0 :*

$$LM_j = \left(\frac{\partial \log L(y, \beta)}{\partial \beta'} \Big|_{\beta = \widehat{\beta}^c} \right)' \widehat{I}^{-1} \left(\frac{\partial \log L(y, \beta)}{\partial \beta'} \Big|_{\beta = \widehat{\beta}^c} \right) \xrightarrow[N \rightarrow \infty]{L} \chi^2(K) \quad (4.45)$$

où $\widehat{\beta}_j$ et $\widehat{\beta}_j^c$ désignent respectivement les estimateurs non contraint et contraint de β_j .

L'estimateur \widehat{I} de la matrice d'information de Fischer peut être obtenu par :

$$\widehat{I} = \sum_{i=1}^N \left(\frac{\partial \log L(y_i, \beta)}{\partial \beta'} \Big|_{\beta = \widehat{\beta}^c} \right) \left(\frac{\partial \log L(y_i, \beta)}{\partial \beta'} \Big|_{\beta = \widehat{\beta}^c} \right)'$$

et où

$$\frac{\partial \log L(y, \beta)}{\partial \beta'} \Big|_{\beta = \widehat{\beta}^c} = \sum_{i=1}^N \frac{\partial \log L(y_i, \beta)}{\partial \beta'} \Big|_{\beta = \widehat{\beta}^c}$$

Pour obtenir la Q statistique et la statistique de test LM, il suffit d'utiliser l'option ARCHTEST dans la procédure AUTOREG.

Figure 4.22: Tests Effets ARCH

```
proc autoreg data=donnees;
  model disp = / archtest dwprob;
  output out=r r=yresid;
run;
```

Le résultat des tests LM et de la Qstat concluent ici clairement au rejet de l'hypothèse nulle d'absence d'effets ARCH et cela quelque soit l'ordre K considéré. Il existe une dépendance temporelle des résidus au carré.

5. Extension des Modèles ARCH / GARCH linéaires

5.1. Application : Value at Risk

cf . Engle R. (2000), "The Use of ARCH/GARCH Models in Applied Econometrics", *Journal of Economic Perspectives*, 15(4), 157-168.

Figure 4.23: Résultats des Tests d'Effets ARCH-GARCH

Q and LM Tests for ARCH Disturbances				
Order	Q	Pr > Q	LM	Pr > LM
1	137.7312	<.0001	137.6889	<.0001
2	253.5812	<.0001	212.9497	<.0001
3	353.0733	<.0001	257.9845	<.0001
4	401.1640	<.0001	266.3164	<.0001
5	473.3338	<.0001	289.1560	<.0001
6	555.3050	<.0001	314.1255	<.0001
7	620.7464	<.0001	326.4233	<.0001
8	678.5805	<.0001	333.8337	<.0001
9	736.3304	<.0001	341.6099	<.0001
10	797.9879	<.0001	351.0227	<.0001
11	852.4882	<.0001	356.4843	<.0001
12	906.1513	<.0001	361.2694	<.0001

cf. site de Philippe Jorion et RiskMetrics <http://www.gsm.uci.edu/~jorion/oc/case.html>

DOSSIER :

- Refaire l'exercice "Orange County Case" de Philippe Jorion à partir des données du cas, en utilisant tous les modèles de volatilité conditionnelle vus en cours.
- Ecrire une synthèse d'une vingtaine de pages en anglais sur les méthodes utilisées pour calculer la VaR.
- Refaire un exercice de validation à la Engle (2000).

5.2. Modèles ARMA-GARCH

C'est Weiss (1986) qui a introduit dans la variance conditionnelle des effets additionnels de variables expliquées. En effet, la modélisation *GARCH* peut être appliquée non au processus initial, mais au processus d'innovation. Ceci permet alors d'introduire divers effets additionnels de variables explicatives soit dans la moyenne conditionnelle, soit dans la variance conditionnelle. Par exemple, on peut considérer un modèle de régression linéaire avec erreurs *GARCH* :

$$y_t = x_t b + \varepsilon_t \quad \varepsilon_t \text{ GARCH}(p, q) \quad (5.1)$$

On peut aussi considérer un modèle *ARMA* avec erreurs *GARCH* :

$$\Phi(L) y_t = \Theta(L) \varepsilon_t \quad \varepsilon_t \text{ GARCH}(p, q) \quad (5.2)$$

Enfin, on peut concevoir un modèle *ARMA* dans lequel la variance non conditionnelle de y peut avoir un effet sur la variance conditionnelle :

$$\Phi(L) y_t = \Theta(L) \varepsilon_t \quad (5.3)$$

$$E\left(\varepsilon_t/\varepsilon_{t-1}\right) = 0 \quad (5.4)$$

$$V\left(\varepsilon_t/\varepsilon_{t-1}\right) = c + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2 + \gamma_0 [E(y_t/y_{t-1})]^2 + \sum_{j=1}^p \gamma_j y_{t-j}^2 \quad (5.5)$$

5.3. Modèles GARCH-M

Engle-Lilien-Robbins (1987) ont proposé des modèles GARCH-M où la variance conditionnelle est une variable explicative de la moyenne conditionnelle. Ces processus semblent ainsi plus adaptés à une description de l'influence de la volatilité sur le rendement des titres.

Definition 5.1. *Un processus y_t satisfait une représentation GARCH-M linéaire en la variance conditionnelle si et seulement :*

$$y_t = x_t b + \delta h_t + \varepsilon_t = x_t b + \delta V\left(\varepsilon_t/\varepsilon_{t-1}\right) + \varepsilon_t \quad (5.6)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ i.i.d. } (0, 1) \quad (5.7)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (5.8)$$

avec $E\left(\varepsilon_t/\varepsilon_{t-1}\right) = 0$ et $V\left(\varepsilon_t/\varepsilon_{t-1}\right) = h_t$.

On peut envisager différentes variantes de la relation entre la variable dépendante y_t et la variance conditionnelle. Par exemple, on peut considérer les cas suivants :

$$y_t = x_t b + \delta h_t + \varepsilon_t \quad \text{Forme Linéaire} \quad (5.9)$$

$$y_t = x_t b + \delta \log(h_t) + \varepsilon_t \quad \text{Forme Log-Linéaire} \quad (5.10)$$

$$y_t = x_t b + \delta \sqrt{h_t} + \varepsilon_t \quad \text{Forme Racine Carrée} \quad (5.11)$$

Toutes ces variantes de GARCH-M peuvent être estimés avec la procédure AUTOREG grâce à l'option MEAN=. Il suffit de préciser :

- MEAN= **LINEAR** pour obtenir une spécification linéaire
- MEAN= **LOG** pour obtenir une spécification log-linéaire
- MEAN= **SQRT** pour obtenir une spécification racine carré

Exemple 1 (Example9.sas) On veut estimer un modèle ARCH(1,1)-M avec une spécification racine carré sur le rendement du SP500, sous l'hypothèse de normalité.

$$y_t = x_t b + \delta \sqrt{h_t} + \varepsilon_t \quad (5.12)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t \text{ N.i.d. } (0, 1) \quad (5.13)$$

Figure 5.1: Estimation d'un Modèle GARCH(1,1)-M sur Rendement SP500

```

proc autoreg data=donnees;
  model d1sp = / garch=( p=1, q=1, mean = sqrt);
  output out=resultsp r=epsilon p=prevision alphacli=0.05
         lcl=lower ucl=upper;
run;

```

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (5.14)$$

Le programme correspondant est le suivant :

Le résultat d'estimation est reporté sur la figure (5.2). On vérifie que SAS nous donne alors l'estimateur du paramètre δ de la spécification GARCH(1,1)-M. On vérifie que plus l'écart type conditionnel augmente, plus le rendement du SP500 augmente toutes choses égales par ailleurs.

Figure 5.2: Estimation d'un Modèle GARCH(1,1)-M sur Rendement SP500

GARCH Estimates					
SSE		0.40117164	Observations		3755
MSE		0.0001068	Uncond Var		.
Log Likelihood		12287.2538	Total R-Square		.
SBC		-24533.353	AIC		-24564.508
Normality Test		3216.2064	Pr > ChiSq		<.0001

Variable	DDL	Estimation	Standard Error	t Value	Approx Pr > t
Intercept	1	-0.000021	0.000425	-0.05	0.9612
ARCH0	1	5.1173E-7	1.0134E-7	5.05	<.0001
ARCH1	1	0.0455	0.003159	14.41	<.0001
GARCH1	1	0.9505	0.003429	277.15	<.0001
DELTA	1	0.0629	0.0514	1.22	0.2208

5.4. Modèles IGARCH

les processus $IGARCH(p, q)$ proposés par Engle et Bollerslev (1986), correspondent au cas d'une racine unitaire dans le processus de variance conditionnelle. Ces modèles sont alors caractérisés par **un effet de persistance dans la variance**. C'est à dire qu'un choc sur la variance conditionnelle actuelle se répercute sur toutes les valeurs futures prévues. L'étude de la stationnarité (au 2nd ordre) d'un processus $GARCH$ revient à démontrer que la variance inconditionnelle est asymptotiquement indépendante du temps (Gourieroux, 1992). Le processus ε_t étant une différence de martingale (admettant des composantes non corrélées de moyenne nulle), on a la propriété suivante :

$$V(\varepsilon_t) = V\left[E\left(\varepsilon_t/\varepsilon_{t-1}\right)\right] + E\left[V\left(\varepsilon_t/\varepsilon_{t-1}\right)\right] = E(h_t) \quad (5.15)$$

Dès lors, la proposition suivante permet de caractériser la notion de stationnarité asymptotique d'un processus GARCH(p, q).

Proposition 5.2. *Un processus ε_t satisfaisant une représentation GARCH(p, q) telle que :*

$$V\left(\varepsilon_t/\varepsilon_{t-1}\right) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (5.16)$$

avec $\alpha_0 \geq 0$, $\alpha_i \geq 0$ pour $i = 1, \dots, q$ et $\beta_j \geq 0$, $j = 1, \dots, p$, est asymptotiquement stationnaire au second ordre si et seulement si :

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1 \quad (5.17)$$

En effet, nous avons vu dans la première section que l'on pouvait obtenir la représentation ARMA suivante sur le processus au carré ε_t^2 .

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + \mu_t - \sum_{i=1}^p \beta_i \mu_{t-i} \quad (5.18)$$

où $\mu_t = \varepsilon_t^2 - V\left(\varepsilon_t/\varepsilon_{t-1}\right) = \varepsilon_t^2 - h_t$ est un processus d'innovation pour ε_t^2 . On peut alors montrer que :

$$E\left(\varepsilon_t^2\right) = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) E\left(\varepsilon_{t-i}^2\right) + E\left(\mu_t\right) - \sum_{i=1}^p \beta_i E\left(\mu_{t-i}\right)$$

Par construction $E\left(\mu_{t-i}\right) = E\left(\varepsilon_{t-i}^2\right) - E\left(h_{t-i}\right) = 0$, puisque nous avons vu que $V\left(\varepsilon_t\right) = E\left(h_t\right)$ dès lors que ε_t est une différence de martingale. On a donc :

$$E\left(\varepsilon_t^2\right) = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) E\left(\varepsilon_{t-i}^2\right)$$

Dès lors, il suffit que les racines du polynôme retard défini par :

$$\Phi(L) = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) L^i$$

soient toutes à l'extérieur du disque unité pour la suite $E\left(\varepsilon_t^2\right) = V\left(\varepsilon_t\right)$ converge. Le processus est alors asymptotiquement stationnaire.

Par conséquent un modèle IGARCH(p, q) est défini par la non stationnarité de son processus de variance conditionnelle :

Definition 5.3. Un processus ε_t satisfait une représentation IGARCH(p, q) si et seulement si

$$V\left(\varepsilon_t/\underline{\varepsilon}_{t-1}\right) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (5.19)$$

avec $\alpha_0 \geq 0$, $\alpha_i \geq 0$ pour $i = 1, \dots, q$ et $\beta_j \geq 0$, $j = 1, \dots, p$ et

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j = 1 \quad (5.20)$$

L'exemple le plus simple est bien évidemment le processus IGARCH(1,1) proposé notamment par Nelson (1990) :

$$V\left(\varepsilon_t/\underline{\varepsilon}_{t-1}\right) = h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad \text{avec } \alpha_1 + \beta_1 = 1$$

Pour ce processus les prévisions de la variance conditionnelles aux différents horizon k sont de la forme :

$$E\left(h_{t+k}/\underline{\varepsilon}_t\right) = (\alpha_1 + \beta_1)^k h_t + \alpha_0 \sum_{i=0}^{k-1} (\alpha_1 + \beta_1)^i$$

Ainsi, lorsque $\alpha_1 + \beta_1 < 1$, le processus ε_t est stationnaire et un choc sur la variance conditionnelle h_t a une influence décroissante et asymptotiquement négligeable sur h_{t+k} quand k tend vers l'infini. Par contre, lorsque $\alpha_1 + \beta_1 = 1$, on a :

$$E\left(h_{t+k}/\underline{\varepsilon}_t\right) = h_t + \alpha_0 k \quad (5.21)$$

En présence d'un terme constant, $E\left(h_{t+k}/\underline{\varepsilon}_t\right)$ diverge avec k . On retrouve exactement les propriétés de prévision sur une marche aléatoire.

On peut préciser ces propriétés en remplaçant ε_t par la quantité $z_t \sqrt{h_t}$ où z_t est un bruit blanc gaussien. Nelson (1990) donne alors une condition de stationnarité au sens strict sur le processus GARCH(p, q).

Proposition 5.4. Le processus (h_t^*, ε_t^*) tel que

$$h_t^* = \alpha_0 \left[1 + \alpha_0 \sum_{k=0}^{\infty} \sum_{i=1}^k (\beta_1 + \alpha_1 z_{t-i}^2) \right] \simeq h_t \quad \varepsilon_t^* = z_t \sqrt{h_t^*} \quad (5.22)$$

est strictement stationnaire si et seulement si :

$$E\left[\text{Log}(\beta_1 + \alpha_1 z_t^2)\right] < 0 \quad (5.23)$$

Enfin, signalons que Ding et Granger (1996,b) établissent que les autocorrélations de la série u_t^2 peuvent être approximées par

$$\rho_k = \frac{1}{3} (1 + 2\alpha_1) (1 + 2\alpha_1^2)^{-\frac{k}{2}} \quad (5.24)$$

et sont donc caractérisées par une décroissance exponentielle typique des processus stationnaires, ce qui est contraire à la conclusion obtenue sur les autocorrélations des processus intégrés en espérance. En d'autres termes, l'observation d'une décroissance lente sur cette fonction d'autocorrélation ne peut être prise en compte par une modélisation IGARCH.

Sous SAS, différentes solutions existent pour estimer un processus IGARCH. La plus simple consiste à utiliser la procédure AUTOREG avec les options GARCH. Celles-ci permettent de préciser un certain nombre de contraintes sur la stationnarité ou non du processus GARCH considéré. Ainsi, on peut spécifier les options suivantes sous la forme **TYPE=** :

- **STATIONARY** : contraint les paramètres du processus GARCH à vérifier les conditions de stationnarité asymptotique.
- **NONNEG** : contraint les paramètres à être non négatifs.
- **INTEGRATED** : permet d'estimer un processus IGARCH(1,1)

Exemple 1 (Example10.sas) : Estimons un processus GARCH(1,1) sur le rendement SP500 en imposant la non-stationnarité avec la procédure AUTOREG :

$$dlsp_t = c + \varepsilon_t \quad (5.25)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t N.i.d. (0, 1) \quad (5.26)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad \alpha_1 + \beta_1 = 1 \quad (5.27)$$

Le programme correspondant est le suivant :

Figure 5.3: Estimation d'un Modèle IGARCH(1,1) sur le Rendement du SP500

```
proc autoreg data=donnees;
  model disp = / garch=( p=1, q=1, type= integrated);
  output out=resultsp r=epsilon p=prevision ;
run;
```

Les résultats sont reproduits sur la figure (5.4). On vérifie bien que la somme des réalisations des estimateurs des paramètres α_1 et β_1 satisfait la contrainte $\alpha_1 + \beta_1 = 1$.

Figure 5.4: Estimation d'un Modèle IGARCH(1,1) sur le Rendement du SP500

Integrated GARCH Estimates					
SSE		0.40124672	Observations		3755
MSE		0.0001069	Uncond Var		.
Log Likelihood		12284.7526	Total R-Square		.
SBC		-24544.813	AIC		-24563.505
Normality Test		3366.7742	Pr > ChiSq		<.0001

Variable	DDL	Estimation	Standard Error	t Value	Approx Pr > t
Intercept	1	0.000461	0.000138	3.35	0.0008
ARCH0	1	3.132E-7	7.9212E-8	3.95	<.0001
ARCH1	1	0.0458	0.003024	15.15	<.0001
GARCH1	1	0.9542	0.002887	330.49	<.0001

6. Modèles ARCH / GARCH asymétriques

La seconde grande approche couvre les modèles **ARCH non linéaires** et plus particulièrement la prise en compte des phénomènes asymétries. L'idée est toute simple : l'effet hétéroscédastique n'est sans doute pas le même suivant que l'erreur précédente est positive ou négative. Deux grandes classes de modèles ont été proposés :

- Nelson (1990) s'est intéressé aux évolutions asymétriques de la variance à l'aide des modèles **EGARCH** (*Exponential Generalized AutoRegressive Conditional Heteroskedastic*).
- Engle et Bollerslev (1986) ont étudié les modèles ARCH à seuils (**TARCH**) où la variance est une fonction linéaire définie par morceaux qui permet différentes fonctions de volatilité selon le signe et la valeur des chocs. Rabemananjara et Zakoian (1991) ont proposé une généralisation avec les modèles les modèles **TGARCH**.

6.1. Modèle EGARCH

Proposé par Nelson (1991), le processus Exponential GARCH ou EGARCH(p,q) donne à la variance conditionnelle la définition suivante :

Definition 6.1. Un processus ε_t satisfait une représentation EGARCH(p,q) si et seulement si :

$$\varepsilon_t = z_t \sqrt{h_t} \quad (6.1)$$

$$\log(h_t) = \alpha_0 + \sum_{i=1}^q \alpha_i g(z_{t-i}) + \sum_{i=1}^p \beta_i \log(h_{t-i}) \quad (6.2)$$

où le résidu normalisé z_t est un bruit faible et où la fonction $g(\cdot)$ vérifie :

$$g(z_{t-i}) = \theta z_{t-i} + \gamma (|z_{t-i}| - E|z_{t-i}|) \quad (6.3)$$

Si l'on pose $a_i = \theta\alpha_i$ et $b_i = \alpha_i\gamma$, la variance conditionnelle de ε_t peut se réécrire sous la forme :

$$\log(h_t) = \alpha_0 + \sum_{i=1}^q a_i z_{t-i} + \sum_{i=1}^q b_i (|z_{t-i}| - E[|z_{t-i}|]) + \sum_{i=1}^p \beta_i \log(h_{t-i}) \quad (6.4)$$

Dans le cas d'un processus EGARCH(1,1), nous avons donc :

$$\log(h_t) = \alpha_0 + a_1 z_{t-1} + b_1 (|z_{t-1}| - E[|z_{t-1}|]) + \beta_1 \log(h_{t-1}) \quad (6.5)$$

Deux remarques doivent être faites à ce niveau :

Remarque 1 *L'écriture porte sur le logarithme de la variance conditionnelle h_t de ε_t , en conséquence aucune restriction n'a besoin d'être imposée sur les différents paramètres de l'équation pour assurer la positivité de h_t .*

Remarque 2 *La variance conditionnelle h_t fait apparaître un effet de signe, correspondant à $a_1 z_{t-1}$, et un effet d'amplitude mesuré par $b_1 (|z_{t-1}| - E[|z_{t-1}|])$.*

Dans ces expressions, la valeur de $E[|z_{t-1}|]$ dépend bien évidemment de la loi supposée de z_t . On a ainsi pour les 4 distributions retenues ici :

$$E[|z_t|] = \sqrt{\frac{2}{\pi}} \quad \text{Loi Gaussienne} \quad (6.6)$$

$$E[|z_t|] = 2 \frac{\Gamma(\frac{v}{2}) \sqrt{v-2}}{\sqrt{\pi} (v-1) \Gamma(\frac{v}{2})} \quad \text{Loi de Student } (v) \quad (6.7)$$

$$E[|z_t|] = \frac{4\xi^2 \Gamma(\frac{1+v}{2}) \sqrt{v-2}}{(\xi + \frac{1}{\xi}) \sqrt{\pi} (v-1) \Gamma(\frac{v}{2})} \quad \text{Loi de Student dissymétrique paramétrée en } \xi \quad (6.8)$$

$$E[|z_t|] = \frac{\Gamma(\frac{2}{\nu})}{\sqrt{\Gamma(\frac{1}{\nu}) \Gamma(\frac{3}{\nu})}} \quad \text{Loi GED de paramètre } \nu \quad (6.9)$$

Pour estimer un modèle EGARCH sous SAS, on peut tout d'abord utiliser la procédure AUTOREG en spécifiant dans la rubrique MODEL la sous option TYPE= EXP dans l'option GARCH pour préciser la nature exponentielle du modèle. Reste que cette procédure n'offre le choix qu'entre deux distributions : la distribution de Student et la distribution normale (par défaut). De la même façon, lorsque l'on précise l'utilisation d'une distribution de Student, SAS ne fournit plus alors d'estimateur du nombre de degrés de libertés v .

Exemple 1 : On souhaite estimer le modèle EGARCH(1,1) suivant sur le rendement du SP500 :

$$dlsp_t = c + \varepsilon_t \quad (6.10)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t N.i.d. (0, 1) \quad (6.11)$$

$$\log(h_t) = \alpha_0 + \alpha_1 g(z_{t-1}) + \beta_1 \log(h_{t-1}) \quad (6.12)$$

$$g(z_{t-1}) = \theta z_{t-1} + \gamma (|z_{t-1}| - E|z_{t-1}|) \quad (6.13)$$

Le programme correspondant avec la procédure AUTOREG est alors le suivant :

Figure 6.1: Estimation d'un modèle EGARCH(1,1) sur le rendement SP500

```
proc autoreg data=donnees;
  model dlsp = / garch=( p=1, q=1,type= exp ) ;
  output out=resultsp r=epsilon p=prevision ;
run;
```

Le résultat est d'estimation reporté sur la figure (6.2). On constate que SAS donne un estimateur des paramètres c, α_0, α_1 et β_1 . On obtient en outre un estimateur du paramètre θ de la fonction $g(\cdot)$, en revanche, on ne dispose pas directement de l'estimateur du paramètre γ mesurant les effets d'amplitude. Ce dernier coefficient est normalisé à 1 dans la phase d'estimation (??).

Figure 6.2: Estimation d'un modèle EGARCH(1,1) sur le rendement SP500

Exponential ARCH Estimates					
SSE		0.40118926	Observations		3755
MSE		0.0001068	Uncond Var		.
Log Likelihood		12347.6656	Total R-Square		.
SBC		-24654.177	AIC		-24685.331
Normality Test		2381.9597	Pr > ChiSq		<.0001
Variable	DDL	Estimation	Standard Error	t Value	Approx Pr > t
Intercept	1	0.000246	0.000124	1.99	0.0471
EARCH0	1	-0.1472	0.0267	-5.51	<.0001
EARCH1	1	0.0982	0.0118	8.35	<.0001
EGARCH1	1	0.9838	0.002887	340.74	<.0001
THETA	1	-0.8423	0.1054	-7.99	<.0001

Une autre façon d'estimer un modèle EGARCH sous SAS, consiste à utiliser la procédure MODELE en écrivant la vraisemblance du modèle où tout au moins de la partie associée à la variance conditionnelle. On rappelle que la fonction ZLAG $n(i, x)$ permet de retourner la valeur retardée de i retards de x , pour un nombre maximum de retards de n , et mets 0 lorsque il existe des valeurs manquantes. Si la valeur d'indice i est omise, la valeur n est considérée. Si n est omis, on considère alors un seul retard.

Exemple 2 : On souhaite estimer le même modèle EGARCH(1,1) que dans l'exemple précédent par la procédure MODEL. Le programme est alors le suivant :

Figure 6.3: Estimation d'un modèle EGARCH(1,1) sur le rendement SP500. Procédure MODEL

```
proc model data = donnees ;
  parms earch0 .1 earch1 .2 egarch1 .75 theta .65 gamma 1.0 ;
  /* Esperance Conditionnelle */
  dlsp = intercept ;
  /* Variance Conditionnelle notée h */
  if (_obs_ = 1 ) then
    h.dlsp = exp(earch0/(1. - egarch1)) ; /*Variance Non Conditionnelle si t=1 */
  else
    h.dlsp = exp(earch0 + earch1*zlag(g) + egarch1*log(zlag(h.dlsp))) ;
    g = theta*(-nresid.dlsp) + gamma*(abs(-nresid.dlsp) - sqrt(2/constant('pi')))) ;
  /* Estimation */
  fit dlsp / fiml method = marquardt ;
run;
```

Dans ce cas, on a introduit le paramètre γ dans les paramètres à estimer, contrairement à la procédure AUTOREG. La variable nresid.dlsp correspond aux résidus standardisés $z_t = \varepsilon_t/\sqrt{h_t}$.

Figure 6.4: Estimation d'un modèle EGARCH(1,1) sur le rendement SP500. Procédure MODEL

Nonlinear FIML Parameter Estimates				
Parameter	Estimation	Approx Std Err	t Value	Approx Pr > t
earch0	-0.14767	0.0271	-5.44	<.0001
earch1	0.202481	0.2387	0.85	0.3964
egarch1	0.983747	0.00295	333.93	<.0001
theta	-0.41093	0.4925	-0.83	0.4041
gamma	0.482969	0.5733	0.84	0.3996
intercept	0.000247	0.000130	1.90	0.0573

Number of Observations		Statistics for System	
Used	3755	Log Likelihood	12348
Missing	1		

6.2. Modèle GJR-GARCH

Afin de prendre en compte la modification d'un coefficient selon la survenue d'un évènement, il est courant dans les travaux économétriques d'introduire une nouvelle explicative construite comme produit d'une indicatrice de l'évènement en question et de la variable initiale. C'est l'idée adoptée par Glosten, Jagannathan et Runkle (1993) qui, partant de l'écriture GARCH(p,q) de base, sont amenés au modèle suivant :

Definition 6.2. Un processus ε_t satisfait une représentation GJR-GARCH(p, q) si et seulement si :

$$\varepsilon_t = z_t \sqrt{h_t} \quad (6.14)$$

$$h_t = \alpha_0 + \sum_{i=1}^q (\alpha_i \varepsilon_{t-i}^2 + \gamma_i \mathbb{I}_{\varepsilon_{t-i} < 0} \varepsilon_{t-i}^2) + \sum_{i=1}^p \beta_i h_{t-i} \quad (6.15)$$

où le résidu normalisé z_t est un bruit faible et $\mathbb{I}_{\varepsilon_{t-i} < 0}$ désigne la fonction indicatrice telle que $\mathbb{I}_{\varepsilon_{t-i} < 0} = 1$ si $\varepsilon_{t-i} < 0$ et $\mathbb{I}_{\varepsilon_{t-i} < 0} = 0$ sinon.

Nous ne considérerons que le cas simple d'un processus GJR-GARCH(1,1) tel que :

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \mathbb{I}_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (6.16)$$

ce que l'on peut encore écrire comme :

$$h_t = \alpha_0 + \alpha_{pos} \mathbb{I}_{\varepsilon_{t-1} \geq 0} \varepsilon_{t-1}^2 + \alpha_{neg} \mathbb{I}_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (6.17)$$

permettant de lire directement les coefficients spécifiques afférents aux résidus positifs $\alpha_{pos} = \alpha_1$ ou négatifs $\alpha_{neg} = \alpha_1 + \gamma_1$.

La seule façon d'estimer un tel processus sous SAS consiste à utiliser la procédure MODEL.

Exemple : On cherche à estimer un modèle GJR-GARCH(1,1) sur le rendement du SP500 tel que :

$$dlsp_t = c + \varepsilon_t \quad (6.18)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t N.i.d. (0, 1) \quad (6.19)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \phi \mathbb{I}_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (6.20)$$

$$\mathbb{I}_{\varepsilon_{t-i} < 0} = \begin{cases} 1 & \text{si } \varepsilon_{t-i} < 0 \\ 0 & \text{si } \varepsilon_{t-i} \geq 0 \end{cases} \quad (6.21)$$

Une autre façon d'écrire la variance conditionnelle est :

$$h_t = \begin{cases} (\alpha_0 + \phi) \varepsilon_{t-1}^2 + \beta_1 h_{t-1} & \text{si } \varepsilon_{t-1} < 0 \\ \alpha_0 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} & \text{si } \varepsilon_{t-1} \geq 0 \end{cases}$$

Le programme correspondant s'écrit alors sous la forme suivante :

Les résultats sont reportés sur la figure (6.6). Ici, on donc :

6.3. Généralisations APARCH et VSGARCH

Introduit par Ding, Granger et Engle (1993) le modèle APARCH est l'un des plus intéressants notamment parce qu'il admet comme cas particuliers plusieurs autres processus existants.

Figure 6.5: Estimation d'un Modèle GJR-GARCH(1,1) sur SP500

```

proc model data = donnees ;

  parms arch0 arch1 garch1 phi ;

  /* Esperance Conditionnelle */
  dlsp = intercept ;

  /* Variance Conditionnelle notée h */
  if zlag(resid.dlsp) > 0 then
    h.dlsp = arch0 + arch1*zlag(resid.dlsp**2) + garch1*zlag(h.dlsp) ;
  else
    h.dlsp = arch0 + arch1*zlag(resid.dlsp**2) + garch1*zlag(h.dlsp) +
      phi*zlag(resid.dlsp**2) ;

  /* Estimation */
  fit dlsp / method = marquardt fiml ;
run ;

```

Definition 6.3. Un processus ε_t satisfait une représentation APARCH(p, q) si et seulement si :

$$\varepsilon_t = z_t \sqrt{h_t} \quad (6.22)$$

$$\sigma_t^\delta = \alpha_0 + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{i=1}^p \beta_i \sigma_{t-i}^\delta \quad (6.23)$$

où $\sigma_t = \sqrt{h_t}$ est l'écart-type conditionnel de ε_t , et z_t est un bruit blanc faible. La positivité σ_t est assurée par les conditions : $\alpha_0 > 0$, $\alpha_i \geq 0$ et $-1 < \gamma_i < 1$, $i = 1, \dots, q$, $\beta_i \geq 0$, $i = 1, \dots, p$, $\delta > 0$.

La stationnarité au second ordre d'un processus APARCH nécessite

$$\sum_{i=1}^q \alpha_i E \left[(|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta \right] + \sum_{i=1}^p \beta_i \sigma_{t-i}^\delta < 1 \quad (6.24)$$

Dans le cas d'un processus APARCH(1,1) on a :

$$\sigma_t^\delta = \alpha_0 + \alpha_1 (|\varepsilon_{t-1}| - \gamma_1 \varepsilon_{t-1})^\delta + \beta_1 \sigma_{t-1}^\delta \quad (6.25)$$

On remarque en particulier que :

- Un processus APARCH(1,1) correspond à une processus ARCH lorsque

$$\delta = 2 \quad \gamma_1 = 0 \quad \beta_1 = 0$$

- Un processus APARCH(1,1) correspond à une processus GARCH lorsque

$$\delta = 2 \quad \gamma_1 = 0$$

Figure 6.6: Estimation d'un Modèle GJR-GARCH(1,1) sur SP500

Nonlinear FIML Parameter Estimates				
Parameter	Estimation	Approx Std Err	t Value	Approx Pr > t
arch0	2.004E-6	3.173E-7	6.32	<.0001
arch1	0.130471	0.0152	8.58	<.0001
garch1	0.909732	0.00967	94.08	<.0001
phi	-0.11873	0.0162	-7.33	<.0001
intercept	0.000319	0.000134	2.39	0.0171

Number of Observations		Statistics for System	
Used	3755	Log Likelihood	12319
Missing	1		

- Un processus APARCH(1,1) correspond à une processus GJR-GARCH lorsque

$$\delta = 2$$

Il est possible de dériver une expression pour le terme $E \left[(|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta \right]$.

Propriété 1 La quantité $E \left[(|\varepsilon_t| - \gamma_i \varepsilon_t)^\delta \right]$ peut s'exprimer en fonction des moments sur $z_t = \varepsilon_t / \sqrt{h_t}$ de la façon suivante :

$$E \left[(|\varepsilon_t| - \gamma_i \varepsilon_t)^\delta \right] = h_t^{\frac{\delta}{2}} E \left[(|z_t| - \gamma_i z_t)^\delta \right]$$

Dès lors, il ne reste plus qu'à évaluer les quantités $E \left[(|z_t| - \gamma_i z_t)^\delta \right]$ pour les différentes distributions que nous pouvons envisager.

Propriété 2 La quantité $E \left[(|z_t| - \gamma_i z_t)^\delta \right]$ associée au bruit blanc faible z_t est définie de la façon suivante. Pour une distribution de Student dissymétrique :

$$E \left[(|z_t| - \gamma_1 z_t)^\delta \right] = \left\{ \xi^{-(1+\delta)} (1 + \gamma_1)^\delta + \xi^{1+\delta} (1 - \gamma_1)^\delta \right\} \frac{\Gamma \left(\frac{\delta+1}{2} \right) \Gamma \left(\frac{v-\delta}{2} \right) (v-2)^{\frac{1+\delta}{2}}}{\left(\xi + \frac{1}{\xi} \right) \sqrt{\pi (v-2)} \Gamma \left(\frac{v}{2} \right)} \quad (6.26)$$

Pour une distribution de Student à v degrés de libertés (cas $\xi = 1$ de la Student dissymétrique) :

$$E \left[(|z_t| - \gamma_1 z_t)^\delta \right] = \left\{ (1 + \gamma_1)^\delta + (1 - \gamma_1)^\delta \right\} \frac{\Gamma \left(\frac{\delta+1}{2} \right) \Gamma \left(\frac{v-\delta}{2} \right) (v-2)^{\frac{1+\delta}{2}}}{1.5 \sqrt{\pi} (v-2) \Gamma \left(\frac{v}{2} \right)} \quad (6.27)$$

Pour une distribution GED de paramètre v :

$$E \left[(|z_t| - \gamma_i z_t)^\delta \right] = \left\{ (1 + \gamma_1)^\delta + (1 - \gamma_1)^\delta \right\} 2^{\frac{\delta-\nu}{\nu}} \frac{\Gamma \left(\frac{\delta+1}{\nu} \right)}{\Gamma \left(\frac{1}{\nu} \right)} \left\{ \frac{\Gamma \left(\frac{1}{\nu} \right) 2^{\frac{-2}{\nu}}}{\Gamma \left(\frac{3}{\nu} \right)} \right\}^{\frac{\delta}{2}} \quad (6.28)$$

Pour une distribution normale (cas GED avec $v = 2$) :

$$E \left[(|z_t| - \gamma_i z_t)^\delta \right] = \left\{ (1 + \gamma_1)^\delta + (1 - \gamma_1)^\delta \right\} 2^{\frac{\delta-2}{v}} \frac{\Gamma\left(\frac{\delta+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} \left\{ \frac{\Gamma\left(\frac{1}{2}\right) 2^{-1}}{\Gamma\left(\frac{3}{2}\right)} \right\}^{\frac{\delta}{2}} \quad (6.29)$$

De la même façon on peut proposer une généralisation des processus GJR-GARCH, avec les modèles VS-GARCH dans lequel la totalité des coefficients peut varier selon le régime et non plus seulement le coefficient du carré de l'innovation passée.

Definition 6.4. Un processus ε_t satisfait une représentation VS-GARCH(1,1) si et seulement si :

$$\varepsilon_t = z_t \sqrt{h_t} \quad (6.30)$$

$$\begin{aligned} h_t = & (\omega_{pos} + \alpha_{pos} \varepsilon_{t-1}^2 + \beta_{pos} h_{t-1}) (1 - \mathbb{I}_{\varepsilon_{t-i} < 0}) \\ & + (\omega_{neg} + \alpha_{neg} \varepsilon_{t-1}^2 + \beta_{neg} h_{t-1}) \mathbb{I}_{\varepsilon_{t-i} < 0} \end{aligned} \quad (6.31)$$

où le résidu normalisé z_t est un bruit faible et $\mathbb{I}_{\varepsilon_{t-i} < 0}$ désigne la fonction indicatrice telle que $\mathbb{I}_{\varepsilon_{t-i} < 0} = 1$ si $\varepsilon_{t-i} < 0$ et $\mathbb{I}_{\varepsilon_{t-i} < 0} = 0$ sinon.

Introduit par Fornari et Mele (1996,1997), il autorise des réponses non symétriques compliquées. Par exemple, des chocs faibles positifs peuvent provoquer une augmentation de la variance conditionnelle supérieure à celle qu'entraînent des chocs faibles négatifs alors que l'inverse se produit pour des chocs élevés selon les valeurs prises par les six coefficients. La variance non conditionnelle de ε_t est donnée par :

$$\sigma_\varepsilon^2 = E(\varepsilon_t^2) = \frac{(\omega_{pos} + \omega_{neg})/2}{1 - (\alpha_{pos} + \alpha_{neg})/2 - (\beta_{pos} + \beta_{neg})/2} \quad (6.32)$$

Exemple : On cherche à estimer un processus VS-GARCH(1,1) sur le rendement SP500 avec la procédure MODEL. Le programme est alors le suivant :

La réalisation de ce même programme sur des données issues d'un modèle simulé avec

$$\begin{aligned} w_{pos} &= 0.1 & \alpha_{pos} &= 0.1 & \beta_{pos} &= 0.5 \\ w_{neg} &= 0.3 & \alpha_{neg} &= 0.5 & \beta_{neg} &= 0.9 \end{aligned}$$

donne les résultats suivants (figure 6.8) :

6.4. Modèles TARCH et TGARCH

Une autre façon de modéliser les asymétries consiste à retenir des modélisations à seuils dans la lignée des modèles TAR (T pour Threshold) de Tong (1990) présentés dans la première section de ce cours. Les modèles TARCH (Zakoian, 1991) et TGARCH (Zakoian, 1994) sont définis de la façon suivante :

Figure 6.7: Estimation d'un Modèle VS-GARCH(1,1) sur Rendement SP500

```

proc model data = donnees ;

  parms arch0neg arch1neg garch1neg arch0pos arch1pos garch1pos ;

  /* Esperance Conditionnelle */
  dlsp = intercept ;

  /* Variance Conditionnelle notée h */
  if zlag(resid.dlsp) > 0 then
    h.dlsp = arch0pos + arch1pos*zlag(resid.dlsp**2) + garch1pos*zlag(h.dlsp) ;
  else
    h.dlsp = arch0neg + arch1neg*zlag(resid.dlsp**2) + garch1neg*zlag(h.dlsp) ;

  /* Estimation */
  fit dlsp / method = marquardt fiml ;

run ;

```

Definition 6.5. *Un processus ε_t satisfait une représentation TGARCH(1,1) si et seulement si :*

$$\varepsilon_t = z_t \sqrt{h_t} \quad (6.33)$$

$$\sqrt{h_t} = \alpha_0 + \alpha_{pos} \mathbb{I}_{\varepsilon_{t-1} \geq 0} \varepsilon_{t-1} + \alpha_{neg} \mathbb{I}_{\varepsilon_{t-1} < 0} \varepsilon_{t-1} + \beta_1 \sqrt{h_{t-1}} \quad (6.34)$$

où le résidu normalisé z_t est un bruit faible et $\mathbb{I}_{\varepsilon_{t-1} < 0}$ désigne la fonction indicatrice telle que $\mathbb{I}_{\varepsilon_{t-1} < 0} = 1$ si $\varepsilon_{t-1} < 0$ et $\mathbb{I}_{\varepsilon_{t-1} < 0} = 0$ sinon.

Ces modèles sont donc similaires aux modèles GJR-ARCH, mais ils spécifient une asymétrie sur l'écart type et non sur la variance conditionnelle.

Exemple : On cherche à estimer un modèle TGARCH(1,1) sur le rendement du SP500 tel que :

$$dlsp_t = c + \varepsilon_t \quad (6.35)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t N.i.d. (0, 1) \quad (6.36)$$

$$\sqrt{h_t} = \alpha_0 + \alpha_{pos} \mathbb{I}_{\varepsilon_{t-1} \geq 0} \varepsilon_{t-1} + \alpha_{neg} \mathbb{I}_{\varepsilon_{t-1} < 0} \varepsilon_{t-1} + \beta_1 \sqrt{h_{t-1}} \quad (6.37)$$

$$\mathbb{I}_{\varepsilon_{t-1} < 0} = \begin{cases} 1 & \text{si } \varepsilon_{t-1} < 0 \\ 0 & \text{si } \varepsilon_{t-1} \geq 0 \end{cases} \quad (6.38)$$

Le programme correspondant est reproduit sur la figure () :

Figure 6.8: Estimation d'un Modèle VS-GARCH(1,1) sur Données Simulées

Nonlinear FIML Parameter Estimates				
Parameter	Estimation	Approx Std Err	t Value	Approx Pr > t
arch0pos	-0.12414	0.0800	-1.55	0.1207
arch1pos	0.47347	0.0570	8.30	<.0001
garch1pos	1.034362	0.0484	21.35	<.0001
arch0neg	0.383686	0.0554	6.92	<.0001
arch1neg	0.227303	0.0268	8.50	<.0001
garch1neg	0.437037	0.0222	19.66	<.0001
intercept	0.51807	0.0392	13.21	<.0001
Number of Observations		Statistics for System		
Used	2000	Log Likelihood	-5002	
Missing	0			

6.5. Modèle QGARCH

Le processus QGARCH (Q pour *Quadratic*) suppose également des asymétries dans la réponse de la volatilité conditionnelle aux innovations. Il a été introduit par Engle et Ng (1993) et Sentana (1995).

Definition 6.6. *Un processus ε_t satisfait une représentation QGARCH(1,1) si et seulement si :*

$$\varepsilon_t = z_t \sqrt{h_t} \quad (6.39)$$

$$h_t = \alpha_0 + \gamma_1 \varepsilon_{t-1} + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (6.40)$$

La variance conditionnelle est donc définie comme une forme quadratique en ε_{t-1} .

Remarque 1 *La forme quadratique $f(\varepsilon_{t-1}) = \gamma_1 \varepsilon_{t-1} + \alpha_1 \varepsilon_{t-1}^2$ étant minimale en*

$$-\frac{\gamma_1}{2\alpha_1}$$

la symétrie de la réponse n'est donc pas obtenue en zéro mais en ce point : à amplitude donnée de l'innovation passée, on a bien un impact sur h_t différent selon le signe de ε_{t-1} .

Par ailleurs, Sentana (1995) montre que les conditions de stationnarité sont identiques à celles dérivées dans le cadre du modèle GARCH, à savoir :

$$\alpha_1 + \beta_1 < 1$$

De plus, comme u est un processus centré, l'expression de son espérance non conditionnelle est également identique à celle obtenue avec un GARCH. En revanche, la kurtosis est croissante avec la valeur absolue de γ_1 , et naturellement égale à celle

Figure 6.9: Estimation d'un Modèle TGARCH(1,1) sur le Rendement SP500

```

proc model data = donnees ;

parms arch0 arch1_pos arch1_neg garch1 ;

/* Esperance Conditionnelle */
dlspl = intercept ;

/* Variance Conditionnelle */
if zlag(resid.dlspl) < 0 then
    h.dlspl = (arch0 + arch1_neg*zlag(resid.dlspl) +
              garch1*zlag(sqrt(h.dlspl)))**2 ;
else
    h.dlspl = (arch0 + arch1_pos*zlag(resid.dlspl) +
              garch1*zlag(sqrt(h.dlspl)))**2 ;

/* estimation */
fit dlspl / method = marquardt fml ;
run ;

```

afférente au GARCH lorsque les deux processus sont confondus, soit pour $\gamma_1 = 0$. Ce gain explique que le QGARCH domine souvent empiriquement le GARCH, ce dernier ayant tendance à sous-estimer l'épaisseur des queues de distribution.

Exemple : On cherche à estimer un modèle QGARCH(1,1) sur le rendement du SP500 tel que :

$$dlspl_t = c + \varepsilon_t \quad (6.41)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad z_t N.i.d. (0, 1) \quad (6.42)$$

$$h_t = \alpha_0 + \gamma_1 \varepsilon_{t-1} + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (6.43)$$

Le programme correspondant est le suivant (figure 6.10) :

Les résultats d'estimation (6.11) confirment que dans le cas du SP500, la présence d'un effet QGARCH, le coefficient γ_1 (noté *gamme*) associé à ε_{t-1} étant largement significatif.

6.6. Modèles LSTGARCH et ANSTGARCH

Avec le processus GJR-GARCH, on modélise deux régimes pour la variance conditionnelle, le choix de l'un ou l'autre étant seulement déterminé par le signe de l'innovation passée. En effet, si $u_{t-1} > 0$ alors $h_t = \alpha_0 + \alpha_{pos} u_{t-1}^2 + \beta_1 h_{t-1}$ et si $u_{t-1} \leq 0$ alors $h_t = \alpha_0 + \alpha_{neg} u_{t-1}^2 + \beta_1 h_{t-1}$. On peut imaginer que le basculement s'effectue de façon moins abrupte et introduire une fonction de transition du type logistique. C'est ce que propose Gonzales-Rivera (1998) avec le Logistic Smooth Transition GARCH qui, en

Figure 6.10: Estimation d'un Modèle QGARCH(1,1) sur Rendement SP500

```

proc model data = donnees ;

  parms arch0 arch1 gamma garch1 ;

  /* Esperance Conditionnelle */
  dlsp = intercept ;

  /* Variance Conditionnelle notée h */
  h.dlsp = arch0 + arch1*zlag(resid.dlsp**2) + garch1*zlag(h.dlsp) +
    gamma*zlag(-resid.dlsp);

  /* Estimation */
  fit dlsp / method = marquardt fiml ;

run ;

```

Figure 6.11: Estimation d'un Modèle QGARCH(1,1) sur Rendement SP500

Nonlinear FIML Parameter Estimates				
Parameter	Estimation	Approx Std Err	t Value	Approx Pr > t
arch0	3.008E-6	4.654E-7	6.46	<.0001
arch1	0.077523	0.00872	8.89	<.0001
gamma	-0.00095	0.000136	-7.01	<.0001
garch1	0.893743	0.0107	83.87	<.0001
intercept	0.000266	0.000135	1.97	0.0494

Number of Observations		Statistics for System	
Used	3755	Log Likelihood	12313
Missing	1		

imposant toujours pour simplifier $p = q = 1$, définit h_t comme

$$h_t = \alpha_0 + \alpha_{pos} u_{t-1}^2 [\Lambda(\theta u_{t-1})] + \alpha_{neg} u_{t-1}^2 [1 - \Lambda(\theta u_{t-1})] + \beta_1 h_{t-1} \quad (6.44)$$

avec

$$\Lambda(\theta u_{t-1}) = \frac{1}{1 + \exp(-\theta u_{t-1})}, \theta > 0 \quad (6.45)$$

Il peut naturellement encore s'écrire

$$h_t = \alpha_0 + (\alpha_{pos} \Lambda(\theta u_{t-1}) + \alpha_{neg} [1 - \Lambda(\theta u_{t-1})]) u_{t-1}^2 + \beta_1 h_{t-1} \quad (6.46)$$

et on voit immédiatement que le coefficient de l'innovation passée est une combinaison linéaire des coefficients afférents à chacun des régimes : lorsque u_{t-1} est grand en valeur absolue et positif il est plutôt proche de α_{pos} , et proche de α_{neg} lorsque u_{t-1} est négatif.

Reprenant l'idée d'une transition douce entre les régimes telle que modélisée par Gonzales-Rivera (1998) pour passer de GJR-GARCH à LSTGARCH, Anderson, Nam et Vahid (1999), définissent à partir du VSGARCH un *Asymmetric Nonlinear Smooth Transition model* selon :

$$h_t = (\omega_{pos} + \alpha_{pos}\varepsilon_{t-1}^2 + \beta_{pos}h_{t-1}) \Lambda(\theta u_{t-1}) + (\omega_{neg} + \alpha_{neg}\varepsilon_{t-1}^2 + \beta_{neg}h_{t-1}) (1 - \Lambda(\theta u_{t-1})) \quad (6.47)$$

où $\Lambda(\cdot)$ est donnée par l'équation (6.45)

Cette écriture permet d'obtenir des réponses non symétriques complexes de la variance conditionnelle aux innovations passées tout en n'imposant pas de changement abrupt de régime.

7. Modèles ARCH et mémoire longue

7.1. Modèle FIGARCH

Lorsque la décroissance exponentielle est trop rapide pour se conformer à celle observée sur la fonction d'autocorrélation, les modèles précédents ne sont pas adaptés. Avec le processus FIGARCH, Baillie, Bollerslev et Mikkelsen (1996) présentent une modélisation qui autorise une décroissance seulement hyperbolique des autocorrélations, et donc a priori intéressante lorsque l'on observe des corrélations encore non nulles pour des ordres élevés⁴. Dans cette présentation, nous limitons les écritures qui suivent aux ordres minimum considérés ici, à savoir $p = q = 1$. Partant de l'équation (??) décrivant le GARCH(1,1) :

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1}$$

soit :

$$(1 - \beta_1 L) h_t = \alpha_0 + \alpha_1 u_{t-1}^2$$

on a :

$$\begin{aligned} h_t &= \frac{\alpha_0}{(1 - \beta_1 L)} + \frac{\alpha_1}{(1 - \beta_1 L)} u_{t-1}^2 \\ &= \frac{\alpha_0}{\beta(1)} + \left[1 - \frac{1 - \delta_1 L}{1 - \beta_1 L} \right] u_t^2 \\ &= \frac{\alpha_0}{\beta(1)} + \vartheta(L) u_t^2 \end{aligned} \quad (7.1)$$

$$\text{avec}^5 \delta_1 = \alpha_1 + \beta_1 \text{ et } \vartheta(L) = \left[1 - \frac{1 - \delta_1 L}{1 - \beta_1 L} \right]$$

Sous le processus IGARCH et la contrainte $\delta_1 = \alpha_1 + \beta_1 = 1$ il vient :

$$\begin{aligned} h_t &= \frac{\alpha_0}{\beta(1)} + \left[1 - \frac{1 - L}{1 - \beta_1 L} \right] u_t^2 \\ \text{soit } \vartheta(L) &= \left[1 - \frac{1}{\beta(L)} (1 - L) \right] \end{aligned} \quad (7.2)$$

le processus FIGARCH considère une puissance fractionnaire sur le terme de différence de cette dernière définition et l'on a donc :

$$\vartheta(L) = \left[1 - \frac{1}{\beta(L)} (1 - L)^d \right], \quad 0 \leq d \leq 1$$

⁴ Comme l'ont montré Ding et Granger (1996,a), la propriété de mémoire longue n'est pas exclusive aux modèles intégrés fractionnaires. Elle est par exemple également vérifiée pour la somme de certains processus AR, pour des modèles à coefficients variables,...

⁵ En effet, soit $\frac{\alpha_0}{1 - \beta_1 L} = \frac{\alpha_0}{\beta(L)} = c$ alors $(1 - \beta_1 L) c = c - \beta_1 c = \alpha_0$ et donc d'une part $c = \frac{\alpha_0}{1 - \beta_1} = \frac{\alpha_0}{\beta(1)}$. Par ailleurs, si on exprime le dernier terme en fonction de u_t^2 et non pas de u_{t-1}^2 : $\frac{\alpha_1}{1 - \beta_1 L} u_{t-1}^2 = \frac{\alpha_1 L}{1 - \beta_1 L} u_t^2$ alors :

$$\frac{\alpha_1 L}{1 - \beta_1 L} = 1 - 1 + \frac{\alpha_1 L}{1 - \beta_1 L} = 1 - \frac{1 - (\alpha_1 + \beta_1)L}{1 - \beta_1 L}.$$

Notons qu'il ne s'agit que d'une facilité d'écriture, en particulier malgré la présence de u_t^2 , on a toujours $E_{t-1}[h_t] = h_t$.

La différence fractionnaire peut être étendue en une série de McLaurin pour donner un polynôme de degré infini en L :

$$\begin{aligned}
 (1-L)^d &= \sum_{i=0}^{\infty} \frac{\Gamma(i-d)}{\Gamma(i+1)\Gamma(-d)} L^i \\
 &= 1 - \sum_{i=1}^{\infty} \frac{d\Gamma(i-d)}{\Gamma(i+1)\Gamma(1-d)} L^i \\
 &= 1 - dL + \frac{d(d-1)L^2}{2!} - \frac{d(d-1)(d-2)L^3}{3!} + \dots
 \end{aligned} \tag{7.3}$$

Dans cette dernière écriture, on a $\sum_{i=1}^{\infty} \frac{d\Gamma(i-d)}{\Gamma(i+1)\Gamma(1-d)} = 1$ ce qui montre bien la parenté existant entre le FIGARCH et le IGARCH. Toutefois, seul le FIGARCH est caractérisé par une décroissance hyperbolique des coefficients de retards et donc par ce que l'on peut nommer une *mémoire longue*. remarquons cependant que DAVIDSON (2002) montre que la mémoire de ce processus augmente lorsque d approche zéro, ce qui est l'opposé de la conclusion obtenue avec les processus fractionnaires en espérance : avec les ARFIMA, la mémoire est une fonction croissante avec d . On doit donc se rappeler que "envisager le modèle FIGARCH comme étant un cas intermédiaire entre un GARCH stable et un IGARCH, de la même façon que l'on pense un processus $I(d)$ sur les niveaux comme étant intermédiaire entre un $I(0)$ et un $I(1)$, est trompeur. Il possède en fait plus de mémoire que chacun de ces processus" (Davidson, 2002, p.8).

7.2. Modèle HYGARCH

Le processus *hyperbolic GARCH* est introduit par Davidson(2002). L'idée est de construire un modèle englobant le FIGARCH afin de tester les restrictions imposées par ce dernier. Il est obtenu en posant⁶

$$\vartheta(L) = 1 - \frac{1}{\beta(L)} \left(1 + \alpha \left((1-L)^d - 1 \right) \right), \quad \alpha \geq 0 \tag{7.4}$$

Selon Davidson, le FIGARCH et le GARCH correspondent respectivement aux cas où $\alpha = 1$ et $\alpha = 0$ ⁷. On note toutefois que lorsque $\alpha = 0$, l'exposant d n'est pas identifiable : la présence de ce terme de *nuisance* affecte la construction des tests d'hypothèses sur α . Il reste toutefois que la proximité de $\hat{\alpha}$ avec l'unité est favorable à la représentation FIGARCH. Sur ce dernier aspect, il convient également de préciser que les propriétés asymptotiques des estimateurs du maximum ou du quasi-maximum de vraisemblance ne sont pas bien assurées⁸. En particulier, même sur grands échantillons, les intervalles de confiance construits selon les règles habituelles (par exemple ± 2 écart-types pour un seuil de confiance de 95%) ne s

⁶ Plus généralement, en posant : $\vartheta(L) = 1 - \frac{\delta(L)}{\beta(L)} \left(1 + \alpha \left((1-L)^d - 1 \right) \right)$

⁷ En fait, lorsque $\alpha = 0$, le HYGARCH correspond à un GARCH pour lequel $\alpha_1 = -\beta_1$ ou encore $\delta_1 = 0$ dans les écritures précédentes.

⁸ Sur ces points, voir Lee et Hansen (1994), Lumsdaine (1996).

7.3. Modèle FAPARCH

Il s'agit d'un processus fractionnaire se caractérisant par une décroissance hyperbolique des autocorrélations mais dans lequel on autorise une asymétrie associée au signe de l'innovation selon le mécanisme caractéristique du processus APARCH. L'équation de volatilité conditionnelle correspondante a donc pour écriture :

$$h_t^{\frac{\delta}{2}} = \frac{\alpha_0}{\beta(1)} + \left[1 - \frac{1}{\beta(L)} (1-L)^d \right] (|u_{t-1}| - \gamma_1 u_{t-1})^\delta \quad (7.5)$$

8. Modèles Multivariés

9. Conclusion

References

- [1] Bera, A. K., et Higgins, M. L. (1993), "ARCH Models: Properties, Estimation and Testing", *Journal of Economic Surveys*, 7(4), 307-366.
- [2] Campbell J.Y., Lo A. et MacKinlay, A.C. (1997), *The Econometrics of Financial Markets*, Princeton University Press.
- [3] Charpentier, A. (2002), *Séries Temporelles : Théorie et Applications*, Polycopié de Cours, Université Paris IX Dauphine.
- [4] Cobbaud, R. (1997), *Théorie Financière*, Economica.
- [5] Colletaz G. "Condhet.src : Estimation de Modèles ARCH sous RATS", *Document de Recherche LEO*, 2002-24.
- [6] Engle, R.F. (1982), "AutoRegressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica*, 50, 987-1008.
- [7] Engle R.F. (2001), "The Use of ARCH/GARCH Models in Applied Econometrics", *Journal of Economic Perspectives*, 15(4), 157-168.
- [8] French, K. (1980), "Stock Returns and the Week-end Effect", *Journal of Financial Economics*, 8(1).
- [9] Gibbons, M.R. et Hess, P. (1981), "Day of the Week Effects and Assets Returns", *Journal of Business*, 54(4).
- [10] Gouriéroux, C. (1992), *Modèles ARCH et Applications Financières*, Collection ENSAE, Economica.
- [11] Granger, C.W.J. et Andersen, A. (1978), *An Introduction to Bilinear Time Series Models*, Vandenhoeck et Ruprecht, Göttingen.
- [12] Keim, D.B. (1983), "Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence", *Journal of Financial Economics*, 12(1).
- [13] Rozeff, M.S. et Kinney, W.R. (1976), "Capital Market Seasonality: The Case of Stock Returns", *Journal of Financial Economics*, 3(4).